



2017

Four Essays on a Student's Expectation that they will Complete College

Martin Gray Hunter

University of Kentucky, gray.hunter@uky.edu

Digital Object Identifier: <https://doi.org/10.13023/ETD.2017.399>

[Right click to open a feedback form in a new tab to let us know how this document benefits you.](#)

Recommended Citation

Hunter, Martin Gray, "Four Essays on a Student's Expectation that they will Complete College" (2017).
Theses and Dissertations--Economics. 33.
https://uknowledge.uky.edu/economics_etds/33

This Doctoral Dissertation is brought to you for free and open access by the Economics at UKnowledge. It has been accepted for inclusion in Theses and Dissertations--Economics by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Martin Gray Hunter, Student

Dr. Chris Bollinger, Major Professor

Dr. Jenny Minier, Director of Graduate Studies

Four Essays on a Student's Expectation that they will Complete College

DISSERTATION

A dissertation submitted in partial
fulfillment of the requirements for
the degree of Doctor of Philosophy
in the College of Business and
Economics at the University of
Kentucky

By
M. Gray Hunter
Lexington, Kentucky

Director: Dr. Chris Bollinger, Professor of Economics
Lexington, Kentucky 2017

Copyright© M. Gray Hunter 2017

ABSTRACT OF DISSERTATION

Four Essays on a Student's Expectation that they will Complete College

It has been common practice in the economics literature to utilize data on observed outcomes and negate what individuals believe or expect will happen in the future. Using responses to a unique set of questions in the 1997 National Longitudinal Survey of Youth (NLSY97) I show that the literature could benefit in several ways by incorporating such data. The leading essay documents a positive association between a student's subjective probabilistic belief that they will complete a four-year college degree and whether or not they attend and complete college. The results indicate the following. First, although overconfident, individuals as young as fifteen are willing and able to answer subjective probabilistic questions concerning education in a cohesive fashion. Second, these expectations are heterogeneous across race, gender, previous academic success, and parent education, and are influential in predicting whether or not they attend and ultimately complete a degree once these characteristics are controlled for. While the magnitude of the effect diminishes when including the standard economic controls, expectations remain significant and play a larger role as the student ages. Parent expectations are also positive and statistically associated with their child's future college success when the student is young but the significance diminishes as the student ages and gathers information related to the costs and benefits of a college degree. These findings indicate that students possess some form of private information that is not being completely captured by the standard variables used by econometricians to predict college attendance and completion.

The second essay uses the NLSY97 to examine how students form and update their college completion expectations as they age out high school. I begin by estimating which factors are utilized by students when forming their expectations while in high school. I find that while these students are taking into account several of the relevant factors associated with college success, they also appear to be neglecting the impact that income and ability have on their likelihood of completing college or are over-relying on poor signals. I then test whether or not students update their expectations in a Bayesian fashion. A Bayesian model is developed. The three ways in which Bayesian students should respond to the acquisition of new information are discussed. Four sources of new information are identified and used in the testing. The testing reveals that students who report either a 0% or 100% chance of completing college do not appear to be Bayesian, but those who report within the 0% and 100% bounds do update in a Bayesian fashion.

The third essay studies the accuracy and alignment of the individual's expectation that they will complete college. I utilize several unique aspects of the NLSY97 to create a measure of alignment based on the predicted probability that the respondent will eventually complete college and their expectation of doing so while either in high school or of college age. I use this measure to answer the following questions. First,

are there any observable differences between those who are aligned and misaligned? Next, do respondents become more aligned as they age and progress out of high school? Last, are those who are more aligned at an early age more likely to reach their outcomes? I find that although the majority of students are overconfident in their belief there are considerable differences in alignment based on several observable characteristics and the availability of information. The alignment of student expectations differ based on parent education, ASVAB percentile, school enrollment, and race. Using two sub-samples of different aged respondents I show that as students age and acquire more information their expectation of completing college becomes more aligned with their estimated probability of completion. I confirm this by examining 700 students who are asked their expectations first in 1997 while in high school then again five years later when they either are in college or the workforce. I conclude by showing that those who are more aligned in either direction with what a model of college completion predicts the more likely they are to eventually reach that outcome.

The final essay examines if the private information contained in the student's expectation that they will complete college is associated with future early career earnings. First I note that there are considerable differences in the frequency of reporting, yearly income, hours worked, and hourly wage for those who predict college success and are successful versus those who do not, as well as those who accurately predict that they will not complete college. I then include these expectations in a wage regression and the estimates suggest that when individuals report their college completion expectations between the ages of 15 and 17 they are not associated with future earnings. However, when asked between the ages of 17 and 22 the reported expectations are positively associated with future wages. There is considerable heterogeneity based on gender, whether they reported at one of the three primary heaping points, and the quantile of the wage distribution in which they were located.

KEYWORDS: Expectations; Education; NLSY97; Bayesian Updating

Author's signature: M. Gray Hunter

Date: September 10, 2017

Four Essays on a Student's Expectation that they will Complete College

By
M. Gray Hunter

Director of Dissertation: Dr. Chris Bollinger

Director of Graduate Studies: Dr. Jenny Minier

Date: September 10, 2017

To my mother, Meems, to whom I am and always will be greatly indebted.

ACKNOWLEDGMENTS

The following dissertation, while an individual work, benefited from the insights and direction of several people. First and foremost, I want to thank my Dissertation Chair, Dr. Chris Bollinger, for his time, help, patience, and encouragement. This dissertation would not have come to fruition without his guidance and valuable advice. I thank my Dissertation Committee members Dr. Tony Creane, Dr. Thomas Ahn, and Dr. Eugenia Toma. I also thank the outside examiner Dr. Jane Jensen. I would also like to thank the participants at the University of Kentucky Brown Bag Seminar, the 2015 and 2016 Kentucky Economic Association, and the 2015 Southern Economic Association for their helpful comments and suggestions. In addition, I thank Dr. Chris Bollinger for admitting me into the program. The financial support I have received from the Department of Economics at the University of Kentucky has helped me in numerous ways. I am also extremely thankful for all the help I have received from Jeannie Graves and Debbie Wheeler.

My family has also played a vital role in my success as both a graduate student and a person and without them this journey would have been impossible. I first thank my mother Mary Reams (Meems) for without her support, sacrifice, and unconditional love throughout my entire life none of this would have been possible. I'm also deeply indebted to my twin sister Allison who consistently moved the bar higher and higher and was supportive of my efforts to keep up with her and all that she has accomplished. Also, to my oldest sisters Dana and Whitney who are also both extremely successful and who have always supported and directed me in all of my pursuits. Last, thanks to my late father Drew who showed me what it meant to truly be a man of honor and virtue and who is largely responsible for shaping me into the man I am today.

In addition, I thank the cohort I was accepted into the program with who provided ongoing support throughout the last five years. I was blessed to be admitted with such a driven, passionate, hardworking, and kind group of individuals and I look forward to hearing of their successes in the future. I thank all of my colleagues in the Economics Ph.D. Program for without their companionship and help my journey through graduate school would have been significantly more strenuous.

TABLE OF CONTENTS

Acknowledgments	iii
Table of Contents	iv
List of Tables	vi
List of Figures	viii
Chapter 1 Introduction to the Dissertation	1
Chapter 2 Literature and Data Overview	7
Chapter 3 Adolescent Econometricians: Smarter than You Might Think. . .	25
3.1 Introduction	25
3.2 Analysis of the NLSY97 Data	27
3.3 Estimation	34
3.4 Results	36
3.5 Sensitivity Analysis	43
3.6 Conclusions	49
Chapter 4 The Formation and Updating of Post-secondary Educational Ex- pectations as Student's Transition from High School	52
4.1 Introduction	52
4.2 Formation of College Completion Expectations	54
4.3 The Change in Students' Expectations from 1997 to 2001	62
4.4 Are Students Bayesian Updaters?	65
4.5 Estimating the Updating Process of Student's Expectations	74
4.6 Conclusions and Future Research	81
Chapter 5 Student Predictions and Alignment of Future College Success . .	84
5.1 Introduction	84
5.2 Unique Aspects of the NLSY97 Data Utilized	86
5.3 Estimation and Results	89
5.4 Conclusions	106
Chapter 6 A Quantile Analysis of Student Expectations of College Comple- tion and Earnings in their Late 20s to Early 30s	110
6.1 Introduction	110
6.2 Data	111
6.3 Analysis of the Wage Data in the NLSY97	115
6.4 Estimation and Results	119
6.5 Conclusions	132

Chapter 7	Conclusion to the Dissertation	134
Appendix A	Dissertation Appendix	142
Bibliography	255
Vita	264

LIST OF TABLES

2.1	NLSY97 Full and Sub-sample Summary Statistics	21
2.2	Summary of the Expectations Sub-samples in the NLSY97	23
3.1	Student Expectations and Outcomes	28
3.2	College Expectations, Attendance, and Completion	29
3.3	College Completion Expectations and Attainment by Race and Gender .	31
3.4	College Completion Expectations by Parent Education	33
3.5	Model and Variable Descriptions	35
3.6	College Expectations, Attendance, and Completion - Probit Model . . .	38
3.7	College Expectations and Completion with 2001 Sample - Probit Model .	41
3.8	College Completion of Overlapping Sample in 1997 and 2001 - Probit Model	42
3.9	Quadratic and Cubic Expectations, Attendance, and Completion	44
3.10	Expectations Indicators, Attendance, and Completion	46
3.11	Parent Expectations and College Achievement	48
4.1	Expectations of Obtaining a Four-year College Degree	55
4.2	Frequency Distribution of Expectations in 1997 and 2001	57
4.3	1997 Expectation Formation and College Completion	59
4.4	Direction and Magnitude of Updated Expectations from 1997 to 2001 . .	64
4.5	1997 and 2001 Expectation Formation	68
4.6	Certainty Priors	71
4.7	Uncertainty Estimates	73
4.8	2001 Expectation Formation - Interior Reporters	76
4.9	Summary Statistics of Bayesian Components	79
4.10	Bayesian Updating Estimates	80
5.1	Full and Sub-sample Summary Statistics	88
5.2	Probit Model Predictions of College Completion - 1997 Sub-sample . . .	90
5.3	Probit Model Predictions of College Completion - 2001 Sub-sample . . .	91
5.4	Summary Statistics for the Secondary Alignment Measure	94
5.5	Estimates of Expectation Alignment in 1997	98
5.6	Estimates of Expectation Alignment in 2001	102
5.7	Expectation Alignment and Completion	104
6.1	NLSY97 Income Summary Statistics	113
6.2	Income, Work, and Prediction Summary Statistics for 1997 Sub-sample .	119
6.3	Income, Work, and Prediction Summary Statistics for 2001 Sub-sample .	119
6.4	Earnings Model Description	120
6.5	Hourly Wage and Expectations Estimates	123
6.6	Hourly Wage and Expectations - Male and Female Estimates	125
6.7	Quantile and OLS Estimates of Hourly Wage and Expectations	128

6.8	Quantile and OLS Estimates of Hourly Wage and Expectations with Indicators	131
-----	--	-----

LIST OF FIGURES

2.1	The NLSY97 Time-line with the Sub-samples Utilized	21
2.2	Frequency Distribution of College Completion Expectations	24
3.1	College Completion Expectations by Highest Grade Completed	30
4.1	Distribution of College Completion Expectations in 1997 and 2001	56
4.2	Distribution of College Completion Expectations for the Overlapping Sub-sample	63
5.1	Primary Alignment Measure for each Sub-sample	92
5.2	Primary Alignment Measure using the Overlapping Sub-sample	93
6.1	Distribution of College Completion Expectations and Income Statistics .	114
6.2	Distribution of Ln(Avg. Hourly Wage) in 2012	116

Chapter 1 Introduction to the Dissertation

This dissertation contributes in several ways to the economics literature by both emphasizing the importance and usefulness of subjective probabilistic data and examining other previously unexplored factors associated with student success in attending and completing college. Although subjective probabilistic data has been utilized in other fields of economics it has not been used in the economics of education realm. Essay one has two primary objectives. The first is to analyze the expectations data reported by high school students to show that individuals as young as fifteen can and do answer subjective probabilistic questions regarding their education in a logical manner. Next, it shows that these data contain private information about the respondent that can be used to predict future success in post-secondary education. To examine the information contained in these expectations I test how they are formed and updated in essay 2; I also test whether or not students update in a Bayesian fashion. Essay three then builds on this by utilizing a model of college completion and examining the accuracy and alignment of the students' expectations. The last essay studies the differences in the earnings distributions based on the reporting and accuracy of the respondents predictions and later shows that the private information in the expectations variable has some predictive power in future hourly wages.

There are several sources of motivation behind this research. First, the rising cost of college, the resulting increase in the level of student debt, and reports of rising unemployment or under-employment by recent college graduates has people reassessing

whether or not college is worthwhile investment.¹ However, what is rarely discussed is the matriculation decision process and elements which contribute to both students' attendance and success at post-secondary institutions. This is even more surprising considering the 6-year graduation rate for first-time, full-time undergraduate students who began seeking a bachelor's degree at a four-year degree-granting institution in the fall of 2009 was 59 percent.²

The second is that obtaining reliable estimates of the returns to education requires that researchers successfully deal with the issues of selection. Manski (1993) outlines a popular approach where researchers model school choice using observed outcome data, assume that students have rational expectations, and that they condition their beliefs on similar variables and process information in a similar way. However, students make schooling choices under uncertainty - uncertainty about personal tastes, individual abilities, and realizations of choice-related outcomes (Zafar, 2011). While the decision process has been examined theoretically (Manski, 1989; Altonji, 1993; Hilmer, 1998; Malumud, 2005) and empirically (Bamberger, 1987; Arcidiacono, 2004; Strange, 2012), most rely on these unrealistic assumptions regarding student expectations. Either implicitly or explicitly, researchers assume that individuals are rational and the formation process is homogeneous. However, if this rational expectations assumption is violated then a number of issues arise. My research uses interpretable subjective data on expectations, as called for by Manski (1993), to show that this rational expectations assumption is flawed.

This research takes advantage of a unique set of questions from the 1997 National

¹While increases in the published cost of college has largely been offset by additional grants or other forms of aid, the out of pocket cost of attending college has increased substantially over the past decade. In 2016-17, the estimated average net tuition and fee price paid by full-time in-state students at public four-year institutions was \$3,770 higher than the net price a decade earlier and \$1,550 higher than the low of \$2,200 in 2009-10. - CollegeBoard (2016); Kroeger et al. (2016); Allen and Seaman (2010); Vedder et al. (2013); Cutler and Lleras-Muney (2006)

²Greenstone and Looney (2012); Department of Education (2013); Shapiro et al. (2015)

Longitudinal Survey of Youth (NLSY97). The survey is nationally representative and collects data on individuals annually from 1997 to 2011, and biannually from 2011 onward. Unlike previous studies, I observe students first while they are in high school and then follow them on an annual basis up until 2013 when they are between the ages of 28 and 33.³ In 1997 and 2001, the survey asked different subsets of students to report their subjective probability of obtaining a four-year college degree by the time they turned 30. In 1997 respondents were in high school and in 2001 most had entered either college or the workforce. To date no other studies have used these data to examine how these expectations relate to post-secondary education or employment outcomes. The probabilistic responses also improve upon past research as other measures of educational expectations that have been used are less precise. These studies have used data where students either reported the highest level of education they expected to complete without any notion of how confident they were in realizing that outcome or they reported the likelihood of completing different levels of education using a Likert scale.

The first essay presents new estimates of the factors that predict a student's decision to attend college and whether or not they complete a four-year degree that economists have not previously explored. Manski (1993) outlined the importance of understanding and identifying how a student's expectations influence their decision to attend college or not. He notes that the estimated returns to education could be flawed when economists rely on the assumption of rational expectations instead of attempting to understand the process by which students use information to form expectations about the costs and benefits of a college degree. Dominitz and Manski (1996) and Manski (2004) both examine these issues, but are limited by small sample

³The NLSY97 followed individuals on an annual basis up until 2011 and then began collecting information every other year. The 2013 data is the most recently released data from the BLS and is utilized in each of the essays below.

sizes and cross-sectional data. Jacob and Wilder (2011) are able to observe students over time and study the role of expectations in degree attainment, but do not examine whether these also influence the individual's decision to attend college. They are also limited by the nature of their data as their measure of expectations is imprecise. This essay adds to the literature by utilizing subjective probabilistic data for students that are followed as they transition from high school into either college or the workforce and observes whether or not they fulfilled their expectations.

The results indicate that, even after controlling for the factors previously identified in the economics of education literature, a student's self-reported probability of completing a four-year college degree plays a role in predicting whether they attended college and if they were successful in obtaining a four-year degree. Next, as students age and progressed through more schooling, their expectations played a larger role in predicting completion. Parent expectations are also estimated to be influential in these models when students are young, but the impact diminishes as the student ages and gathers information related to the costs and benefits of a college degree. The estimates signal that students' expectations contain a form of private information that is helpful in determining post-secondary schooling attendance and success that isn't being captured by the previously identified controls. These findings suggests that youth and econometricians may possess and utilize different data when analyzing schooling decisions.

The second essay begins by examining whether or not students incorporate the factors that have been identified with college success when forming their expectations. I estimate which factors students utilize when forming their educational expectations using two different aged sub-sets of the data. I find that while high school aged students take into account several of the relevant factors, they also appear to be neglecting the impact that income and ability have on their likelihood of completing college. Overall, students are extremely overconfident in their ability to complete

a four-year college degree and it is possible that they either are not incorporating these factors into their expectations. I then test whether or not students update their expectations as a Bayesian model predicts. A Bayesian model is presented and I discuss the three ways in which students should respond to the acquisition of new information. I then present four new sources of information that become available between 1997 and 2001 to a subset of respondents who reported their expectations in both periods. I show that while students do appear to be utilizing the new information when updating their expectations, they do not appear to be true Bayesian updaters. I hypothesize two possibilities as to why this is not the case.

Essay 3 examines the accuracy of the students' predictions that they will complete a four-year college degree. The goal is to identify if the accuracy of predictions vary based on observable characteristics and if so how this information can be used to help students form their expectations more accurately. This could serve as a way to increase the graduation rates at post-secondary universities and reduce the percentage of those who take on excessive student loans and drop out. I am able to use the large representative sample of the NLSY97 to compute the probability that each student will complete a four-year degree by 2013 based on characteristics identified in the previous literature. Two new measures of alignment are then created that use this predicted probability and the student's expectation. I use these measures to test whether alignment differs based on observables, improves as the respondent ages, and if it is associated with their realized outcomes. I find that on average, student alignment differs based on age, parent education, ASVAB percentile, school enrollment, and race. As they age and acquire more information, their expectation of completing college becomes more aligned with their estimated probability of completion. Also, those whose expectations were similar to their predicted probability of completing college were more likely to eventually reach their expected outcome.

The last essay includes these educational expectations in a wage regression to test

whether the private information contained within them has some predictive power for future hourly earnings. I begin by identifying differences in the frequency of reporting, yearly income, and hours worked for those who predict college success and were successful versus those who were not, as well as those who accurately predict that they would not complete college. I find that those who successfully predicted that they would complete college reported receiving any income during 2012 at a much higher rate than those who failed to predict success and those who predicted they would not complete college. They also reported a higher annual income and worked nearly 300 more hours per year than both groups. Those who failed to predict completing college reported receiving any income at a higher rate, a higher income, and that they worked more hours than those who predicted they wouldn't complete college, but the differences were marginal. The results from the wage regression indicate that when individuals report these expectations between the ages of 15 and 17, they are not associated with future earnings. However, when asked between the ages of 17 and 22 the reported expectations are positively associated with future wages. There are considerable differences based on gender, whether they reported at one of the three primary heaping points, and which quantile of the wage distribution they were located.

Chapter 2 Literature and Data Overview

Education and Expectations in the Literature

Since Manski's (1993) discussion of students as adolescent econometricians, economists have recognized the importance of identifying the information that students use when deciding whether or not to pursue a post-secondary education. Students face similar issues to econometricians in that they are attempting to predict their future success using the data available to them. However, students may possess different data, have different knowledge of the economy, and may process information in different ways than econometricians. He focuses on the universal assumption that students form expectations homogeneously.¹ Using a model of school choice where students either do or do not condition based upon their ability, he shows how the assumption of a homogeneous expectations formation process for all students significantly alters the predictions from the model and leads to two identification problems.² This research does not specifically focus on these identification issues, but reviews the literature on school choice and identifies the factors that have been shown to predict college success. I then test whether student expectations play a role that econometricians have not previously found, examine how respondents update their educational expectations as they age, and identify certain characteristics that make students more accurate predictors.

The education literature contains useful research on college choice and educational

¹Economic studies of schooling behavior use this both implicitly and explicitly in the majority of previous research but rely on different information processing rules and hypothesized conditioning variables.

²The first is that, not knowing how youth perceive the returns to schooling, researchers cannot infer their decision processes based on their schooling choices. Second, is that not knowing a youth's decision processes, the objective returns to schooling cannot be estimated from data on realized outcomes.

expectations. Hossler and Gallagher (1987) first reviews the past literature on how student background characteristics, aspirations, and achievements interact with student expectations of college, and student access to the availability of higher education institutions. They then use this literature as motivation for developing a three-phase model of student college choice and how policy makers might utilize it. Phase one, the predisposition phase, is simply students deciding they want to attend college. Stage and Hossler (1989) examine this phase using high school students from Indiana. They show difference in college expectations for students conditional on their parents' aspirations, family characteristics, and gender. They do not follow students over time and use Likert scale data for expectations which limits their analysis.³ In phase two the potential matriculants seek more information about colleges and universities, during which they discuss how a well-researched choice does not assure a rational choice. Following the search phase, the students enter the third phase of college enrollment where they decide which school to attend. Hossler et al. (1999) later discuss each of the three separate phases but only include a discussion of college aspirations in terms of how they relate to the initial choice of whether or not to search for information on college opportunities. They do not incorporate them in terms of how students search for information or if they play a role in ultimately attending college. They also do not examine whether or not these expectations play a role in matriculation for students who are enrolled in a post-secondary institution. MacLeod (2008) discusses the potential role of college expectations in each of the three phases for students but do not empirically test their hypothesis. I improve upon these studies in several ways.⁴

While college choice, expectations, and aspirations have been studied heavily in

³Several of the models in essay two follow Stage and Hossler (1989)'s framework, but I expand upon them by testing more factors of formation and examine how students' expectations change after they made their college choice.

⁴Also see Hossler and Stage (1992); John et al. (1996); St. John et al. (2005) for more on research on education, expectations, and aspirations in the education literature.

the education literature, the topic has not been as widely studied by economists. A large portion of the economics of education literature examines how the cost of attendance influences the attendance decision of students. Hilmer (1998) presents a theoretical model based on school pricing and ability that outlines the decision that students make when deciding to attend college or not, and if so, whether they should attend a two or four-year degree granting institution initially. The model predicts that when the net cost of attendance decreases, more students should attempt to obtain a college degree; the opposite is true for when costs rise. Dynarski (1999) follows this and uses social security benefits to estimate that a \$1,000 increase in student aid resulted in a 3.6 percentage point increase in attendance, and students completed an additional 1/10th a year of post-secondary schooling. Manski and Wise (1983), Leslie and Brinkman (1988), and Dynarski (2000) find similar results using a variety of different semi-natural experiments and program implementations.

Others have examined the role of gender, race, socioeconomic status, and information in predicting college attendance and completion. Perna (2000) discusses how during the 1990s the percentage of undergraduates that were either African American or Hispanic increased significantly, yet they remained underrepresented in both the undergraduate and bachelor degree recipient populations compared to their relative representation in the traditional college-aged population. Her findings suggest that financial aid in the form of loans reduced the probability of enrolling for African Americans compared to other races. The probability of enrolling also increased for Whites and African Americans as the local unemployment rate increased, but Hispanics were uninfluenced. Last, she identified several other non-cost related factors that have heterogeneous effects on college attendance based on race.

Bailey and Dynarski (2011) later demonstrate that overall college completion rates have increased, but the gaps in college entry, persistence, and graduation for children from high and low income families have increased. They note this was especially

true for women, driven by an increase in the education of daughters of high-income parents. Hoxby and Avery (2012) and Hoxby et al. (2013) both find that low-income high achieving students either did not attend college or attended two-year institutions at a much higher rate than their high-income high achieving counterparts due to a lack of information about the "net" cost of college.⁵ Their experiment supplied students with information regarding available scholarships and estimates of the out-of-pocket cost of several surrounding colleges for income and achievement similar students. After supplying students with this information, they discovered that their attendance and completion rates increased such that they were comparable to those of similar ability. Similar to their research, there are also federal outreach and student services programs, referred to commonly as TRiOs, in the United States that are designed to identify and provide services to individuals from disadvantaged backgrounds.⁶ While information such as this is difficult to measure, it is possible that it could be contained in a student's expectations of completing college in the future. Those from low income or disadvantaged backgrounds might benefit the most from additional information, but the research below does not limit the analysis based on these characteristics as a representative study of all those considering a post-secondary education is more expandable to the entire population. Future research will limit the analysis based on these characteristics.

Economic research examining educational expectations and student outcomes dates back to the 1960s with Willam H. Sewell, Otis D. Duncan, and their colleagues being credited with the seminal work demonstrating a positive correlation between expectations and both educational and occupational attainment (Sewell et al., 1970).

⁵Also see Terenzini et al. (2001)

⁶TRIO includes eight programs targeted to serve and assist low-income individuals, first-generation college students, and individuals with disabilities to progress through the academic pipeline from middle school to post-baccalaureate programs. See (U.S. Department of Education, 2017) for more information on these programs.

A hurdle that the expectations literature was forced to overcome was the apprehension that economists had regarding the accuracy and usefulness of self-reported subjective data.⁷ To combat this, Dominitz and Manski (1996) designed and applied an interactive computer-administered personal interview (CAPI) survey that elicited expectations of high school students and undergraduates regarding their income if they were to complete different levels of schooling. They found that respondents, even as young as high school juniors, were willing and able to respond meaningfully to questions eliciting their earnings expectations in probabilistic form. Dominitz and Manski (1997) followed by collecting data on the one-year-ahead income expectations of members of American households in a survey of economic expectations and found again that self-reported expectations were reliable and useful when making predictions.

The following papers have utilized expectations data and echo the findings of Dominitz and Manski (1996,1997) that expectations data contain useful information. Bernheim (1990) used the Social Security Administration's Retirement History Survey to examine the elderly's expectations of Social Security benefits during a preretirement period and how they responded to new information. He found that the expectations of benefits were relatively noisy, but they could be used to make accurate predictions about the individual's received benefits in the future. He also notes sizable differences in accuracy conditional on race, gender, and marital status. Smith et al. (2001) used four waves of the Health and Retirement Survey (HRS) to test whether longevity expectations were related to the observed mortality of individuals. They discovered that subjective beliefs about longevity were consistent with individuals' observed survival patterns, and that deaths were signaled through lower longevity expectations reported in previous periods. Hurd and Mc-

⁷Machlup (1946) is cited as the initial critique of using expectations data in economic models by Jacob and Wilder (2011).

Garry (2002a) used the HRS to confirm Smith et al. (2001)’s findings and expanded upon them by examining how respondents modified their survival probabilities in response to new information such as the onset of a new disease or condition. They found that respondents modified their survival probabilities in response to new information, and that the subjective survival probabilities could be used to predict survival; those who survived in the panel reported survival probabilities approximately 50% greater at baseline than those who died. Lochner (2007) uses the NLSY97 to empirically examine belief updating of the perceived probability of arrest and its criminal deterrence effects. His estimates suggest that a higher perceived probability of arrest could reduce criminal participation. In each case, the authors found that individuals’ expectations were reliable and had predictive power. This study adds to this literature by focusing on the educational expectations of teens and young adults and how they relate to educational outcomes and future wages.

Fischhoff et al. (2000) analyzes how teens in the first round of the NLSY97 responded to the 18 questions about the probability of future life event occurring. While the objective of the study was to examine the risk perceptions of teens, the authors comment extensively on the accuracy of the responses when compared to public health figures concerning the mentioned event. The most relevant for my research are those related to the educational expectations of students. They examined the percent chance a student believed they: would be in school next year, obtain a high school diploma by age 20, and earn a four-year college degree by age 30.⁸ Their findings indicate that as the time from anticipated completion increased, so did student error. When predicting whether they’d be a student next year, the mean was nearly identical to public health statistics for children of a similar age at the time. As students looked further into the future, they were overly optimistic about their chances

⁸The high school expectations variable is not examined in this research as graduating high school is a more common outcome.

of graduating both high school and college, reporting means of nearly 93% and 73% respectively, when in reality the graduation rates from both were approximately 84% and 25% respectively.

Other findings from Fischhoff et al. (2000) include differences in reporting patterns between those still in school and those no longer enrolled. Also, a strong correlation between expectations of graduating both high school and college was found with the reported percent of their peers that planned to go to college. Overall, they found that for the majority of categories the students', and parents', expectations were relatively accurate, at least on average, when compared to national public health data. This indicates that both adults and teens as young as fourteen are willing and able to answer probabilistic questions and are relatively accurate.⁹ While the findings are interesting, they are only able to compare the students' expectations to the outcomes of similar students during that time. This study isn't limited by this as both the students' expectations and their outcomes as of their late 20s and early 30s are observed. This allows me to compare their expectations to their eventual outcomes and examine the predictive power they may have.

To this point, only a few studies in the economics literature have directly inspected the role of student expectations in post-secondary achievement.¹⁰ Zafar (2013) uses a unique panel data set of 161 Northwestern University sophomores which includes the students' subjective expectations about several of their future academic outcomes.¹¹ He incorporates these expectations in a choice model where selection of a college major is made under uncertainty in an attempt to explain the markedly different choices

⁹Some expectations questions were asked to those who were 14, but they do not include the educational expectation questions.

¹⁰Kao and Tienda (1998) use the National Education Longitudinal Study of 1988 (NELS:88) to analyze how educational aspirations are formed and maintained from eighth to twelfth grades among a single cohort of youth, but does not follow the students as the progress into college.

¹¹These include the percent chance that: they graduate with a GPA of 3.5 or better, enjoy courses, hours spent on coursework per week, find a job, enjoy work, work flexible hours, etc....

by males and females. He found that students were overconfident about their future academic performance, but their expectations were useful when making predictions. Students also updated these expectations rationally. However, the use of students from Northwestern results in selection issues that the author cannot combat.¹² Students who attend and complete their first year at Northwestern possibly differ in both the information they receive and how they process and utilize it than a more representative sample of students across different colleges. My research has limited issues of selection as it uses a representative sample from across the United States. It also observes students prior to, during, and after they have completed college.¹³

Jacob and Wilder (2011) utilize a combination of data from Monitoring the Future, High School and Beyond, the National Educational Longitudinal Survey of 1988, and the Educational Longitudinal Survey of 2002 to answer a number of questions related to educational expectations. They show that high school seniors expecting to complete at least some college rose from the mid-1970s up until the late 1990s where they peaked around 93%. However, while expectations rose during this period actual attainment did not follow suit, thus the predictive power of expectations fell. They found that over sixty percent of students updated their educational expectations at least once between 8th grade and eight years post high-school, and that it is primarily based upon the acquisition of new information regarding their ability through signals like GPA, test scores, or graduating. They also note that students who did not update as significantly, those who they consider more aligned at a young age, generally had a higher likelihood of fulfilling their expectations. Rather than defining alignment

¹²Northwestern is one of the top universities in the United States, it only admits approximately 13% of applicants, has an ACT composite range of 31-34, and an SAT composite range of 1390-1560. Forbes lists Northwestern University as a top 15 college overall. It's ranked as the #14 in private colleges, #9 in research universities, and #2 overall in the Midwest. (Forbes, 2016)

¹³Arcidiacono et al. (2012) has a similar study with Duke students that focuses primarily on major choice and the subjective expectations of income associated with each. The selection issues present for the Northwestern students are similar for this study as Duke is also a premier University.

based on how the respondent's expectations changed I introduce two new measures of alignment that incorporate the predicted probability that the student would complete college. My results from these new measures confirm their findings that those who were more aligned early on were more likely to reach their predicted outcome.

The literature above highlights the importance of examining and incorporating subjective probability data into models of school choice. They show that self-reported expectations have predictive power in terms of a number of different outcomes. If student expectations can also be used to predict college attendance and completion it would reaffirm Manski's notion that students face similar issues to econometricians, but may possess different data, have different knowledge of the economy, and may process information in different ways when it comes to school choice. Several relevant factors have been identified as influencing both a student's decision to attend a post-secondary institution and whether or not they are able to graduate. I base the controls used in the models below on these findings. Jacob and Wilder (2011)'s analysis is similar to my research, but they are limited by how students reported their expectation of attending and completing college. Students reported their highest expected level of attainment without a measure of confidence attached to it which limited them to examining large, discontinuous differences in education.¹⁴ Another contribution to this literature is the presentation and analysis of the probabilistic data available in the NLSY97 as they have not been utilized widely in the economics literature. I include a discussion of the data in the section below.

Data: The 1997 National Longitudinal Survey of Youth

The motivation behind the 1997 National Longitudinal Survey of Youth (NLSY97) was to provide researchers the opportunity to identify characteristics that define the

¹⁴They also utilize some expectations reported on Likert scale, but it is not the primary focus of their analysis.

transition from schooling to the labor market for youth during this period. It was designed such that participants would be representative of the United States in 1997 for those born during the years 1980 through 1984.¹⁵ They first administered surveys in 1997 to 8,984 individuals who were born between January 1, 1980 and December 31, 1984 and thus ranged from 12 to 17 years old at the date they were initially interviewed. Following the initial 1997 survey, the youth were interviewed subsequently on an annual basis up until 2011, after that the survey was and will be administered biennially. This study utilizes data from 1997 up to the first biennial responses from 2013. Overall, approximately 80% of those initially surveyed in 1997 also completed the most recent 2013 survey used.¹⁶

In the first round, both the selected youth and one of their parents each participated in a one hour long personal interview. The youth questionnaire focused in detail on their schooling and employment activities, but also included questions regarding their family background, social behavior, health status, and the expectation of future events occurring. Detailed information was also collected on a number of other topics, including: education, training, and achievement scores; employment; household characteristics; income, assets, and program participation; health; crime and substance abuse; and attitudes, expectations, non-cognitive ability measures, and general activities. The inclusion of the parent is beneficial in that it provided more extensive and reliable information regarding the youth's family background, household environment, and general history.

In addition to the round 1 interview, the youth were also incentivized to take a computer-adaptive version of the Armed Services Vocational Aptitude Battery

¹⁵Bureau of Labor Statistics, U.S. Department of Labor (2013)

¹⁶The incentive to participate in the survey has varied in several ways. Monetarily, the respondents typically received between \$10 and \$20 for participation and it depended on which round they responded to. The survey administrators offered different levels of incentives to respondents in an effort to study the effects of incentive level on survey participation throughout.

(ASVAB) at a separate time. The ASVAB is composed of 12 subject tests that are meant to measure knowledge and skill of participants in a number of key areas including mathematics, science, and language. The measure used throughout this research is a summary measure created by the NLS staff from the following four subtests: Mathematical Knowledge, Arithmetic Reasoning, Word Knowledge, and Paragraph Comprehension.¹⁷ This is used over the individual test scores as these four tests are more likely to be predictive of academic success than the other eight which appear to be more related to the skills associated with success at a trade school.¹⁸ The respondents received a different ASVAB score via mail several months after the exam. This score was for several different categories such as math and verbal and was accompanied by pamphlet with information about each.¹⁹

The exam was given to individuals in groups of five to ten under standardized conditions at more than 280 sites across the country. The tests were primarily administered at Sylvan Learning Centers, but some temporary testing sites were also established at hotels, community centers, and libraries when a testing center was not available.²⁰ Respondents were given \$75 to participate and in the end 7,712 or 79.3%

¹⁷Although the formula is similar to the AFQT score generated by the Department of Defense for the NLSY79 cohort, this variable reflects work done by NLS program staff and is neither generated nor endorsed by the Department of Defense. For a detailed description on how the measure was created visit: <https://www.nlsinfo.org/content/cohorts/nlsy97/topical-guide/education/administration-cat-asvab-0>

¹⁸The other 8 tests were: Assembling objects, Auto information, Coding speed, Electronics information, General science, Mechanical comprehension, Numerical operations, and Shop information.

¹⁹A sample of what students received in the mail can be found in the appendix. Respondents also took an "Interest Finder" examine which was more related to their strengths and interests in a career sense which they also received information on.

²⁰Testing was conducted by the Department of Defense (DOD) and took place from the summer of 1997 through the spring of 1998; the DOD used the NLSY97 participants as part of a larger effort to establish new norms for the test, which is primarily used for military enlistment screening.

of the sample took the computerized ASVAB.²¹ In an attempt to maximize sample size, those that did not take the ASVAB were given the mean score of their gender and ethnic group and an imputation indicator variable was included in all regressions where the ASVAB percentile was used.²² While the ASVAB is generally used as a proxy for ability, I utilize it as a source of information for the students and use the students grades from 8th grade and high school as measures of academic ability.²³

Students' high school transcript data was collected by the BLS directly from the schools in two waves for those who provided written permission. These data can be broken down into individual terms for which the student was in school and contains their courses taken, grades, and school related activities in which they were involved. These data are available for 6,232 or 69% of the individuals who participated in the first round of the study. The low reporting level could signal possible selection by either parents or students who chose to release their grades. However, students were asked in the survey to self-report their grades from both 8th grade and high school and the response rate was much higher than the school collected grades. While Black et al. (2011) found measurement error issues for individuals who reported their

²¹This is considerably lower than the 93.9% completion rate for the pencil and paper format of the comparable Armed Forces Qualifying Test given to the National Longitudinal Survey of Youth 1979 participants. Per Donna Rothstein and Mark Loewenstein at the BLS, this is likely the result of a number of factors. First, the NLSY97 test was given in the first year of the survey compared to the second year for the NLSY79. Second, \$50 was offered to participants in 1979 compared to \$75 in 1997; \$50 in 1979 would have been worth approximately \$117 in 1997. Third, a Spanish-language version might not have been available which would have hurt the NLSY97 participants more as it is composed of approximately 21% Hispanics compared to just 15% for the NLSY79. Also, in the NLSY79, special needs students and non-English speakers were excluded which was not the case for the NLSY97. The data supports this as the non-takers were primarily Hispanic.

²²Specifications that also excluded all of those whose score was imputed were also estimated and noted if differences occurred.

²³Neal and Johnson (1996) argue that it is an adequate measure of human capital for students that have yet to enter the labor market. They found that its' exclusion lead to biased estimates when examining student outcomes. They focus on wages when the students are 26 to 29, but endogenous selection into schooling is discussed. However, Bollinger (2003) and Black et al. (2011) find that measurement error in the AFQT and ASVAB, when used as proxy for human capital, can lead to biased estimates of coefficients for correctly measured variables that are correlated with human capital.

highest level of education attainment, when comparing the transcript grades to the students' self-reported grades they were relatively precise. It is possible that those who selected into providing access to their transcript GPA differed in their reporting from those who did not, and as such models utilizing them were estimated separately. The models that utilized this data in the research below suggest a selection issue, and thus to increase sample size and avoid this issue those that used the self-reported grades are emphasized.²⁴

The NLSY97 is composed of 2 subgroups. The first, which is referred to as the cross-sectional sample, consists of 6,748 respondents that are representative of those born in United States during the period of January 1, 1980 to December 31, 1984. The second, which is referred to as the supplemental or over-sample, includes 2,236 respondents that are either Hispanic/Latino or African American born during the same period. The full sample is 51% male and 49% female, and 52% Non-black/non-Hispanic (categorized as white), 26% Black non-Hispanic, 21% Hispanic or Latino, and 1% Mixed or Multiracial.²⁵ The majority of the models below utilize the full sample and include an indicator for whether or not the individual was in the representative portion.

This research relies heavily on two subsets of the NLSY97. The first includes only those who were 15 or 16 years old as of December 31st, 1996. These 3,450 respondents were asked to report their educational expectations during round 1 in 1997. The second subset includes 1,902 individuals who were asked to report their educational expectations during round 5 in 2001; no age restrictions were placed on those who were asked the questions in 2001 so it contains a larger age range of participants than the

²⁴For the estimates that include the transcript data email gray.hunter@uky.edu.

²⁵The mixed race individuals were classified as Black in this study per the suggestion of Dr. Chris Bollinger.

1997 subset.²⁶ Approximately 700 individuals were asked to report their expectations in both 1997 and 2001. Figure 2.1 below presents a time-line and overview of the full sample and each of the sub-samples utilized. The summary statistics for each follow in table 2.1. The 1997 and 2001 subsets have a similar gender and race composition compared to the full sample. Each are nearly 51% male and 49% female, and are approximately 52% white. The 1997 subset is 27% black while the 2001 subset is 25.4% black. The two are also 20.8% and 23% Hispanic respectively. When they are separated by cross-sectional and over-sample categories, there is more variation from the entire sample. The 700 overlapping students are of 49% male and 51% female, and 50.8% white, 25.4% black, and 23.7% Hispanic. A means test for both race and gender show that the two expectation sub-samples are not statically different from the full sample, and the 700 overlapping differs only in that it slightly over-samples Hispanics.²⁷ The next section discusses the 1997 and 2001 sub-samples that were asked to report their educational expectations.

²⁶This group was a part of an experiment where a fraction of the individuals were randomly chosen to answer the expectations questions in each appropriate topical section, rather than all together in one section as in the 1997 survey. This was done to examine the difference in response rates as a result of the different grouping and ordering. I have been unable to locate the results of this study.

²⁷Household income was estimated for approximately 27% of the sample due to a lack of reporting. The mean income based on race, parent education, and number of parents in the household was given to those who did not report and an indicator for imputed household income was included in each estimation. This was done only for essays 1-3, it was not imputed for those in essay 4 where it is used as the dependent variable.

Figure 2.1: The NLSY97 Time-line with the Sub-samples Utilized

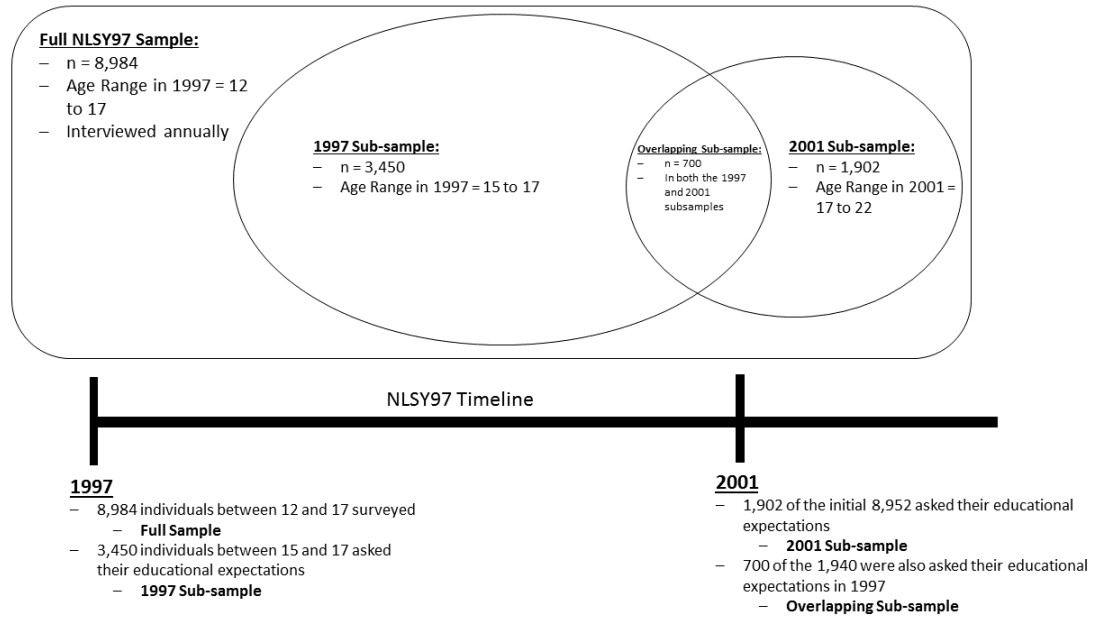


Table 2.1: NLSY97 Full and Sub-sample Summary Statistics

	Sample Type	Individuals	Mean Age	Male	White	Black	Hispanic
Full Sample	Cross-Section	6,614	13.92	51.39%	69.44%	16.84%	13.71%
	Over-Sample	2,162	14.02	51.25%	0%	55.64%	44.36%
	Total	8,776	13.98	51.36%	52.34%	26.40%	21.26%
1997 Sub-sample	Cross-Section	2,582	15.47	50.39%	69.71%	16.77%	13.52%
	Over-Sample	868	15.47	51.84%	0%	57.72%	42.28%
	Total	3,450	15.47	50.75%	52.17%	27.08%	20.75%
2001 Sub-sample	Cross-Section	1,440	18.90	51.67%	68.13%	15.90%	15.97%
	Over-Sample	462	18.96	49.13%	0%	55.19%	44.81%
	Total	1,902	18.92	51.05%	51.58%	25.44%	22.98%
Overlapping 700	Cross-Section	517	15.52	49.32%	68.86%	14.89%	16.25%
	Over-Sample	183	15.42	49.18%	0%	55.19%	44.81%
	Total	700	15.49	49.29%	50.86%	25.43%	23.71%

Notes: Mean age for the full sample, the 1997 sub-sample, and the overlapping 700 is as of December 31st, 1996; the mean age for the 2001 sub-sample is the age at the time of interview. The overlapping 700 are the 700 respondents who answered the expectations questions in both 1997 and 2001.

Expectations in the NLSY97

A unique set of questions included in the NLSY97 are those that pertain to the respondent's expectation that a future event will occur. The expectations questions range from things as simple as the percent chance the individual believes that they will eat pizza next year to more complicated issues such as the percent chance they go to jail or die in the following year. While a number of other surveys elicit respondents' expectations, the NLSY97 is unique in how it asks individuals to report them. Most surveys ask individuals to respond to the likelihood of an event occurring using a Likert type scale where participants choose from a range of possibilities similar to "never" or "least likely" to "always" or "very likely". Others simply ask respondents to report their highest expected level of educational attainment without any measure of confidence. Examples of this include the Educational Longitudinal Study of 2002, the High School and Beyond, and the National Educational Longitudinal Study of 1988. The benefit of the NLSY97 is that individuals are asked to report the percent chance between 0 and 100 that they believe the specified event will occur. This elicits the student's subjective probability that the event will occur which allows for a more detailed analysis.

This study focuses on the respondents' educational attainment expectations, more specifically the belief that they will have completed a four-year college degree by the time they are 30 years old. The survey asked the two subsets of the sample in 1997 and 2001 the following question, *"Now think ahead to when you turn 30 years old. What is the percent chance that you will have a four-year college degree by the time you turn 30?"* Participants then responded with their belief on the scale from 0 to 100 percent.²⁸ Summary statistics for each sub-sample responded are included below followed by Figure 2.2 which presents the frequency distribution of these expectations

²⁸Bureau of Labor Statistics, U.S. Department of Labor (2013)

for both the 1997 and 2001 respondents.

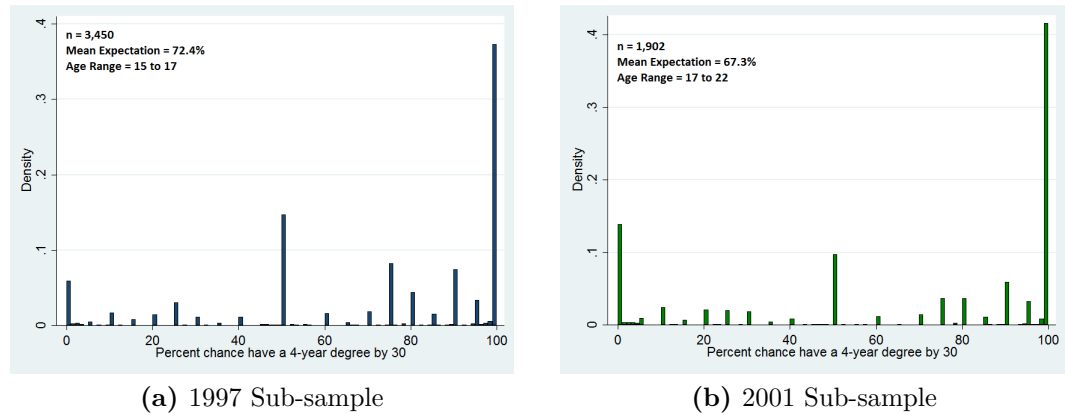
Table 2.2: Summary of the Expectations Sub-samples in the NLSY97

	N	Mean	Standard Deviation	Minimum	Maximum
1997 Subsample	3,450	72.4%	31.43	0	100
2001 Subsample	1,904	67.3%	38.48	0	100
Overlapping Sample					
In 1997	700	74.29%	30.81	0	100
In 2001	700	64.94%	39.88	0	100
Δ Expectations	700	-9.35	40.1	-100	100

Notes: Statistics reflect the answer to the following question, "Now think ahead to when you turn 30 years old. What is the percent chance that you will have a four-year college degree by the time you turn 30?". The standard deviation is percentage points and not percent.

The 1997 cohort has a higher mean, similar median, and less variation in their expectations than the 2001 cohort. This could signify that students are updating as they age and progress through schooling, and they are becoming more certain about the outcome whether it be that they will or will not obtain the degree. This idea is reinforced when examining the 700 individuals who reported their expectations in both 1997 and 2001. For these individuals, the average change in expectations was a decrease of 9.35 percentage points, with a standard deviation of 40.1. On average, students were overly optimistic about their chance of obtaining a four-year college degree. By 2013, the majority of individuals in both sub-samples were between 29 and 33 years old. For the younger, 1997 cohort the average of all of the students' expectations is 72.4% while only 25% of them went on to obtain a four-year degree by 2013. For the 2001 cohort, which is on average two years older than the 1997 cohort, the statistics are 67.3% and 27% respectively.

Figure 2.2: Frequency Distribution of College Completion Expectations



From the distributions above we see that there is considerable heaping around 0, 50, 75, 90, and 100 percent with the largest share of respondents reporting a 100 percent chance of obtaining a degree. In 1997, 5.9% of the sample reported a 0 percent chance of completing a four-year degree, 14.7% reported a 50 percent chance, 8.2% believed they had a 75 percent chance, 7.4% believed they had a 90 percent chance, and 35.4% reported a 100 percent chance. In 2001, 13.8% reported a 0 percent chance, 9.7% reported a 50 percent chance, 3.7% reported a 75 percent chance, 5.9% believed they had a 90% chance, and 38.5% reported a 100 percent chance. Heaping is an issues and is addressed within each of the subsequent essays. This section is meant to give an overview of the NLSY97 and the primary data utilized in the essays below. The specific components and sub-samples utilized in each essay are discussed separately within each.

Chapter 3 Adolescent Econometricians: Smarter than You Might Think.

3.1 Introduction

The recent increase in the out-of-pocket cost of attending college has led to a number of discussions, including how to reduce these costs and whether a degree is still a worthwhile investment. However, the matriculation decision process and elements which contribute to both students' attendance and success are seldom discussed. I present new estimates of the factors that predict a student's decision to attend college and whether or not they complete a four-year degree that economists have not previously explored. I focus on the predictive power of students' self-reported probability of obtaining a four-year college degree. Later I consider how the predictive power of these expectations vary as the student ages and progresses through different levels of schooling. We anticipate that as students gain more information about the costs and benefits of the degree as well as their ability to complete it, their expectations should play a larger role in predicting their outcomes.

Manski (1993) outlined the importance of understanding and identifying how a student's expectations influence their decision to attend college or not. He notes that the estimated returns to education could be flawed when economists rely on the assumption of rational expectations instead of attempting to understand the process by which students use information to form expectations about the costs and benefits of a college degree. Dominitz and Manski (1996) and Manski (2004) both examine these issues, but are limited by small sample sizes and cross-sectional data. Jacob and Wilder (2011) are able to observe students over time and study the role of expectations in degree attainment, but do not examine whether these also influence the individual's decision to attend college. They are also limited by the nature of their data as their measure of expectations is imprecise. This study improves upon these by utilizing

data that follows students as they transition from high school into either college or the workforce and observes whether or not they obtained a four-year college degree by the time they turn thirty.

A unique set of questions from the National Longitudinal Study of 1997 (NLSY97) allows for an examination of how a student's subjective probability of completing a college degree changes as they progress through schooling and what role it plays in their outcomes. The survey asks a subset of students to report their probability of obtaining a four-year college degree by the time they turn 30, once in high school and again when some have entered either college or the workforce. To date no other studies have used these data to examine the role of expectations in educational outcomes. The probability scale is an improvement over other studies as most use data where students respond on a Likert scale regarding their educational expectations, or are simply asked the highest level of education they expect to complete without any notion of how confident they are in realizing that outcome.

The results indicate that, even after controlling for the factors previously identified in the literature, a student's self-reported probability of completing a four-year college degree plays a role in predicting whether they attempt to complete a college degree and if they are successful in obtaining a four-year degree. As students age and progress through more schooling, their expectations play a larger role in predicting completion. Also, parent expectations are estimated to be influential when students are young, but the impact diminishes as the student ages and gathers information related to the costs and benefits of a college degree. The estimates signal that students' expectations contain a form of private information that is helpful in determining post-secondary schooling attendance and success that isn't being captured by the previously identified controls, and suggests that youth and econometricians may possess and utilize different data when analyzing schooling decisions. While the combination of interpretable subjective data on expectations and student choice data

could help solve the identification problem discussed by Manski (1993), further analysis is needed to discover how students are forming and updating these expectations.

3.2 Analysis of the NLSY97 Data

Prior to discussing how expectations are related to several of the previously identified factors associated with post-secondary achievement I first show that, although students are either extremely overconfident or aspirational in their ability, their expectations are still directly related to educational outcomes. First, I calculate the correlation coefficient for: expectations and attendance for the 1997 cohort, and expectations and completion for both cohorts.¹ For attendance, the estimated coefficient is .4028 signaling a weak, yet positive association. For the same cohort, the correlation of .3320 between expectations and completion is smaller but still positive. When the older 2001 cohort is used this correlation increases to .4587. This is to be expected as the majority of these students are of college age and should have a better understanding of what it takes to complete a college degree.

Similar to Hurd and McGarry (2002b), I also compare the expectations of those who go on to complete the outcome of interest to those who do not. If a relationship is present, we expect that those who report lower(higher) expectations initially are less(more) likely to complete any college and a four-year degree. The results are in the table below. I divide the 2001 sub-sample into 3 categories based on their highest level of academic achievement up to that point in order to further analyze the relationship. In each case, those that go on to both attend and complete college have higher initial expectations on average than those who do not. The students in the 1997 sample who go on to attend college had an average expectation of 84.3% while those who did not had an average of 58.5%. For completion, the 877 who go on to

¹I do not calculate the correlation coefficient for attendance and expectations for the 2001 cohort as the majority of this sample is already college aged.

complete the degree had an average initial expectation of 90.49%, while the 2,573 who did not had an average of 66.17%. A similar pattern is present in the 2001 sample. Also, columns 2 and 4 suggest a sense of false optimism for those in 1997 sample who do not go on to attend or complete college. When they are compared to those who have completed 12th grade or less from the 2001 cohort, they have noticeably higher expectations. This subset of the 2001 sub-sample is on average a year and half older than the 1997 sample when reporting their expectations, so it is possible that during this time the students who had high expectations but a low probability of being able to complete the degree are receiving a signal that they use to update their expectations downward.

Table 3.1: Student Expectations and Outcomes

Sample	Student Outcomes in 2013			
	Attend College	Does Not Attend College	Completes College	Does Not Complete College
1997 Expectations	84.30% (1,850)	58.53% (1,600)	90.49% (877)	66.17% (2,573)
2001 Expectations				
- \leq 11th Grade	85.59% (333)	41.47% (459)	93.01% (144)	53.38% (532)
- = 12th Grade	83.49% (292)	43.14% (360)	95.95% (120)	52.69% (648)
- $>$ 12th Grade	88.36% (461)	n/a	97.79% (258)	76.93% (200)

Notes: Mean expectations are reported for those associated with each outcome. The number of students in each category is in parenthesis. r below represents the calculated correlation coefficients.

- $r_{attend,exp_97} = .4028$

- $r_{complete,exp_97} = .3320$, $r_{complete,exp_01} = .4587$

Table 3.2 divides the expectations for each cohort into eight bins and presents a corresponding measures of matriculation, completion, and persistence. The number of and percent of students in each cohort that attempt and complete a four-year degree are included, as well as a measure of persistence that reports the fraction of students who completed at least one year of college who went on to obtain a four-year degree.

This is used as a rough measure of persistence and is included in the last column of the table. Excluding those who reported a 100% chance of obtaining a four-year degree, there is a distinct positive relationship between reported expectations and both attempting to obtain a college degree and completion of that degree for both the 1997 and 2001 cohorts; this relationship is also present for expectations and persistence outside of the 0% and 100% bins for both cohorts. The differences in attendance and completion rates between the 91 to 99% and the 100% bins suggest possible issues with overconfidence, misreporting, or a lack of understanding of probability. These issues are examined in the sensitivity analysis section.

Table 3.2: College Expectations, Attendance, and Completion

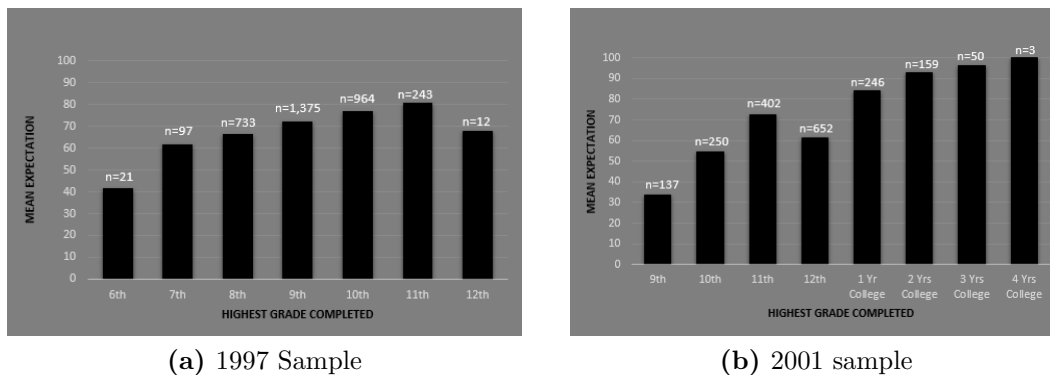
	Percent Chance Have a 4 Year Degree by 30	Number of Students	Number that Attempt Degree	Number that Complete Degree	Percent that Attempt Degree	Percent that Complete Degree	Percent that Start and Complete Degree
1997 Sub-sample	0%	205	37	9	18%	4.4%	24.3%
	1 - 25 %	278	58	5	20.9%	1.8%	8.6%
	26 - 49 %	103	34	3	33%	2.9%	8.8%
	50 %	508	170	37	33.5%	7.3%	21.8%
	51 - 75 %	433	243	92	56.1%	21.2%	37.9%
	76 - 90 %	476	317	149	66.6%	31.3%	47.0%
	91 - 99 %	226	174	112	77.0%	49.6%	64.4%
	100 %	1,221	852	470	69.8%	38.5%	55.2%
	Total	3,450	1,885	877	54.6%	25.4%	46.5%
2001 Sub-sample	0 %	263	39	3	14.8%	1.1%	7.7%
	1 - 25 %	179	41	1	22.9%	0.6%	2.4%
	26 - 49 %	64	12	1	18.8%	1.6%	8.3%
	50 %	184	54	6	29.3%	3.3%	11.1%
	51 - 75 %	126	70	18	55.6%	14.3%	25.7%
	76 - 90 %	211	126	48	59.7%	22.7%	38.1%
	91 - 99 %	143	119	85	83.2%	59.4%	71.4%
	100 %	732	628	360	85.8%	49.2%	57.3%
	Total	1,902	1,089	522	57.3%	27.4%	47.9%

Notes: Number that Attempt Degree includes those who reported their highest grade completed as 1 year of college or more.

Now that a relationship between expectations and achievement has been identified, it is useful to examine how expectations vary conditional on some of the factors identified in the previous literature also found to be related to post-secondary schooling achievement. First I examine differences based on the highest level of completed schooling when asked about their expectations. Eighty-nine percent of students in the 1997 cohort had completed between 8th and 10th grade when asked their expectations, while for the 2001 cohort ninety percent had completed between 10th grade

and 2 years of college.² Figure 3.1 displays how the students' expectations vary based upon the highest reported level of schooling completed when interviewed.

Figure 3.1: College Completion Expectations by Highest Grade Completed



The figures show that expectations of attaining a college degree rise as students complete more education with the exception of 12th grade, or a high school degree. This seems to be the pivotal point for students in regards to their belief about obtaining a college degree. This is unsurprising as it is the point where many students decide whether or not to pursue a college degree or other activities. However, if students are weighing this option in their initial expectations formation, then such a large discontinuity should not be present. It's plausible that students are realizing that the costs, both monetary and psychic, of obtaining a college degree are higher than anticipated. Another possibility is that they discover that the benefits of a high school degree are higher than they believed, which would serve to diminish the believed benefit of a college degree in comparison. This is controlled for and briefly examined in the empirical section below.³ Those who report having completed college at least four years of college in 2001 are not utilized in the models.

²There is more variation within the 2001 cohort as the respondents were selected from the entire sample.

³There are a myriad of other factors that could also play a role in this revision process that will be explored in future work.

Examining the expectations and completion of the 1997 cohort conditional on gender and race, we find that women have higher expectations than men by roughly 8 percentage points and also go on to complete college 30% of the time compared to just 21% for men. Whites and blacks have similar expectations, yet whites on average are nearly twice as likely to graduate college as blacks. Whites and blacks both have higher expectations and completion rates than do Hispanics. This pattern is also present when comparing across both race and gender. Overall, women have higher expectations and are more likely to graduate college than men of the same race. These relative patterns are also present for the 2001 cohort. Expectations appear to be heterogeneous across race and gender, but are positively related to eventual attainment. The expectations and completion statistics for both cohorts used are presented based on gender, race, and then race and gender in the table below.

Table 3.3: College Completion Expectations and Attainment by Race and Gender

		N	Mean	Median	Std. Deviation	Percent that Complete College
1997 Sub-sample	Male	1,751	68.14%	75%	32.4	20.6%
	Female	1,699	76.69%	94%	30.72	28.5%
	White	1,800	74.42%	90%	31.40	32.3%
	Black	934	73.42%	90%	32.24	18.0%
	Hispanic	716	66.45%	75%	32.11	13.3%
	White Male	918	68.72%	80%	32.54	28.1%
	Black Male	452	70.85%	80%	32.28	14.6%
	Hispanic Male	381	63.54%	70%	32.23	9.5%
	White Female	882	79.80%	95%	29.13	36.7%
	Black Female	482	75.83%	95%	32.06	21.2%
	Hispanic Female	335	69.75%	75%	31.70	17.3%
	Total	3,450	72.35%	85%	31.91	24.5%
2001 Sub-sample	Male	971	63.16%	80%	39.08	22.2%
	Female	931	71.68%	95%	37.45	30.3%
	White	981	69.92%	90%	38.15	37.1%
	Black	484	67.24%	90%	38.27	16.5%
	Hispanic	437	61.62%	75%	39.10	12.4%
	White Male	510	65.69%	85%	38.60	31.8%
	Black Male	240	62.83%	80%	39.70	12.5%
	Hispanic Male	221	57.69%	60%	39.12	10.9%
	White Female	471	74.51%	99%	37.15	42.9%
	Black Female	244	71.57%	95%	36.37	20.5%
	Hispanic Female	216	65.64%	85%	38.76	13.9%
	Total	1,902	67.33%	90%	38.52	26.18%

Notes: Expectations are reported between 0 and 100%, so the standard deviation are percentage point changes rather than percent.

Table 3.4 presents similar statistics based upon the parent education level for the students. A strictly positive relationship exists between students' college expectations and the amount of education their parents have as well as for completion. Students whose parents both have a high school degree or less not only have the lowest expectations and completion rates, but they also have the most variation in their beliefs. As the education level for students' parents increases, expectations also increase and less variation is present. For example, in the 1997 cohort the standard deviation of expectations for students' whose parents both have a college degree is approximately half that of students whose parents hold a high school degree or less (18.44 vs 34.35 percentage points). This difference is larger for the older, 2001 cohort. Parents who have experienced college can provide their children with certain insights as to the costs and benefits of a degree that a student whose parents have not completed college would not be able to. This could be in the form of making sure their children are prepared academically, what to expect the first year on campus, or even through helping them on assignments. Thus, students with more educated parents should have access to more information and therefore have more aligned expectations. The pattern could also be that they are less influenced by the monetary costs of college, as their parents more than likely have higher incomes and more certain they can fund their education. Further testing is needed to disentangle the effects.

Table 3.4: College Completion Expectations by Parent Education

	Parent's Degree	N	Mean	Median	Std. Deviation	Percent that Complete College
1997	Both HS or Less	1843	64.28%	75%	34.35	12.0%
	At Least 1 has Some College	1040	78.77%	90%	27.93	30.5%
	Both Have Only Some College	177	81.44%	90%	24.78	33.9%
	At Least 1 has a College Degree	810	86.42%	98%	22.05	52.5%
	Both Have at Least College Degree	303	90.25%	100%	18.44	67.7%
2001	Both HS or Less	970	56.27%	58.5%	39.99	12.4%
	At Least 1 has Some College	622	74.26%	98%	35.92	31.7%
	Both Have Only Some College	129	79.77%	100%	31.38	31.8%
	At Least 1 has a College Degree	453	87.70%	100%	25.16	57.2%
	Both Have at Least College Degree	181	94.33%	100%	14.80	72.9%

Notes: Education level is for genetic parents as of 1997. Does not take into account if both of the parents are in the household or not when interviewed.

Last, I briefly discuss the relationship between the ASVAB, students' expectations, and post-secondary outcomes. The data reveal that a higher ASVAB percentile is associated with both higher expectations and a higher likelihood of both attempting and completing a college degree. For example, those in the 1997 sample that are in the lowest ASVAB quantile have an average expectation of 62%, and while 30% of the this sample completes at least 1 year of college, only 6% go on to complete the degree. As you move up the quantiles, expectations and successful post-secondary outcomes increase accordingly.⁴ This pattern is consistent for the 2001 sample. This roughly signals that students understand and incorporate the information tested by the ASVAB into their expectations, both of which are positively related to post-secondary schooling success. There are a number of other factors that go into each and further analysis is needed.⁵

Analyzing the data, it is clear that at least on average students' expectations re-

⁴Per Donna Rothstein, the students received their results several months after taking the exam. The results included only their performance on each of the subject tests and did not include the percentile used in this data.

⁵Tables comparing the ASVAB percentile to both expectations and achievement for both samples can be found in the appendix which can be obtained by emailing the author at gray.hunter@uky.edu.

garding the attainment of a four-year college degree are heterogeneous based on the highest level of previous academic achievement, gender, race, parent education, and ASVAB percentile. Yet these expectations also appear to be related with students' post-secondary educational outcomes. On average, those with higher expectations of obtaining a college degree are associated with a higher probability of both attending college and ultimately graduating with a four-year degree. However, it is plausible that after controlling for these, schooling success measures, and other factors identified by the literature that expectations will simply contain the information already included in these controls rather than be another source of information that can be used to predict post-secondary schooling success. The next section tests this hypothesis.

3.3 Estimation

I first examine college matriculation or attendance. Next, the completion of a four-year degree is analyzed. For attendance, students are given a 1 if they report completing at least 1 year of college by 2013 and 0 otherwise. This was done instead of using any credits received as reporting of credits in the NLSY97 is significantly worse than years of education.⁶ For completion, those who report having obtained a Bachelor's degree or higher by 2013 are given a 1 while those who report less are given a 0.⁷ The estimates of both matriculation and completion will be via maximum likelihood using a probit model. The estimated marginal effects at the means will be reported in the tables and discussed below.⁸

Each model uses a set of covariates that have been previously identified in the

⁶If anything, this should bias the estimated association of expectations on attendance downward.

⁷The NLSY97 has a variable for the highest grade and degree reported ever. So, those who reported completing some college or completing a Bachelor's degree in previous years that did not respond to the 2013 survey are still included.

⁸OLS results were similar and can be found in the appendix.

economic literature, many of which are included in the literature review above, as potentially influencing or predicting the outcomes of interest. These include demographics; parent education levels; location; family characteristics; self-reported grades in 8th grade and/or high school; the students' peers' plans for college and the level of involvement of their high school teachers.⁹ The specifics of each category are outlined in the table below.

Table 3.5: Model and Variable Descriptions

<u>Dependent Variables:</u>	i. Student completes at least 1 year of college ii. Student completes a 4 year college degree	
<u>Variable of Interest:</u>	The percent chance the student believes that they will obtain a 4 year college degree by the time they are 30	
<u>Control Variables:</u>	Demographics - Male - White - Black	Parent Education - One parent has some college - Both parents have some college - One parent has a college degree - Both parents have a college degree
	Family - Only child - Household members under 18 - Mother, Father, or both parents absent - Income quantile	Schooling: - Self-reported grades in 8th grade (levels) - Self-reported grades in HS if applicable (levels) - ASVAB Percentile - Enrollment in high school indicator
	Location - Urban - MSA Central City - MSA Non Central City - 3 Census Tracts (W, NE, NC)	Other: - % of peers that plan to go to college - Teacher quality and involvement indicators

Notes: For each categorical variable, that which is not listed is used as the base case. For grade controls, if a student is still in high school then the 8th grade controls are used, when they have graduated or left school the high school grade controls are used. In each case the student reported their grades in levels. Transcript data for high school GPA is available but low reporting severely reduces sample size. Estimates from specifications that have included it do not differ drastically. The teacher quality and involvement controls are reported in levels by the student. The student reports whether they believe their teachers overall were 'good' and whether they were 'interested in the students success.' They responded with strongly agree, agree, disagree, or strongly disagree.

Since there are two distinct samples, estimates are obtained using each separately. Recall, the 1997 cohort is composed of 15 to 17 year old students in grades ranging from 8th to 12th. Parents of these students are also asked to report the expectations of their child's probability of completing a college degree. These will be discussed in the later section. The 2001 sample contains some students who are in high school,

⁹See Kiuru et al. (2007) for the role of peer groups in adolescents' educational expectations and adjustment.

college, or neither. Therefore, they will be used only in the completion models and indicators for enrollment in high school, a two-year college, and a four-year college will be included.¹⁰ Approximately 700 students reported their college completion expectation in both round 1 and round 5. I estimate the completion model first using their expectations in 1997 and then for 2001 to compare how the predictive power of expectations change as students age and transition out of high school.

3.4 Results

For both outcomes, estimates from three difference specifications are presented. The first includes only expectations on the right hand side. The next does not include expectations, but only the controls mentioned previously. The last combines the two above such that both expectations and the other controls are all included. The first two specifications are used to show that expectations and the previously identified controls by themselves have explanatory power. When the two are combined we observed how the coefficient on expectations changes and if the statistical significance is altered. Note, that while the NLSY97 lends itself to a panel analysis it is not being used as such for this paper. The subscript t is used to indicate the two different subsets of the data being used, and the subscript n denotes the time added until the student's outcome is observed in 2013. These estimates suggest that expectations are highly related to both outcomes even when controlling for the various factors associated with college attainment, however they do not imply that a causal relationship exists as they are endogenous and a suitable instrument has not yet been identified.¹¹ The models and the corresponding estimates are discussed below.

¹⁰Specifications for the attendance model using those who reported still being enrolled in high school in 2001 were estimated. The results can be found in the appendix.

¹¹Jacob and Wilder (2011) present several reasons why this is the case.

Matriculation

The dependent variable for matriculation is an indicator for whether the student ever completed at least one year of college.¹² The probit model below is used for the estimation where Φ represents the cumulative distribution function of the standard normal distribution and X is a vector of controls that were previously mentioned. Since none of the respondents in the 1997 sample had completed high school at the time of their interview only their grades from 8th grade are used.

$$Prob(Attend College_{i,1997+n}) = \Phi(Expectations_{i,1997}, X'_{i,1997}\beta) \quad (3.1)$$

The related matriculation estimates for the 1997 sample are included in columns 1-3 in the table below. Expectations are found to be a positive and statistically significant factor in predicting the probability that a student attends college. The addition of the control variables decreases the estimated impact by approximately 47% which was expected due to its correlation with several of the other controls. However, it still suggests that, at the means, a 10 percentage point increase in expectations increases the probability that a student completes at least 1 year of college by 3.7 percentage points on average, *ceteris paribus*. I consider this economically relevant as decreasing the out-of-pocket costs for a student by \$1,000 is found to have a similar impact in the literature.¹³ The addition of expectations also increases the percent of variation explained by the model from 26.1 to 28.6%. For simplicity, the other estimated coefficients have been suppressed but each had the hypothesized influence

¹²Ideally, this would indicate whether students simply attended college since a large portion of students that drop out do so after one semester or less. Unfortunately, reporting of individual college credits was incomplete for the majority of students so it could not be utilized.

¹³See Manski and Wise (1983), Leslie and Brinkman (1988), Dynarski (1999), and Dynarski (2000).

excluding one of the race controls.¹⁴ Estimates from the full model that includes all of the controls can be found in the appendix.¹⁵

Table 3.6: College Expectations, Attendance, and Completion - Probit Model

	Y = 1 if completed at least 1 year of college			Y = 1 if completed a 4 year college degree		
	(1)	(2)	(3)	(4)	(5)	(6)
College Expectations_97	0.00699*** (0.000312)		0.00373*** (0.000350)	0.00589*** (0.000300)		0.00225*** (0.000303)
Demographic Controls	No	Yes	Yes	No	Yes	Yes
Parent Education Controls	No	Yes	Yes	No	Yes	Yes
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade & ASVAB	No	Yes	Yes	No	Yes	Yes
Peer and Teacher Controls	No	Yes	Yes	No	Yes	Yes
Observations	3450	3403	3403	3450	3366	3366
Pseudo R^2	0.125	0.261	0.286	0.123	0.326	0.340

Notes: Demographic controls include male, white, and black. Parent education include indicators for if one or both of the respondents' parents have some college or a four-year degree. Location controls include rural, 1 of 4 census tracks, and MSA. Family controls include number of members in the household under 18, whether student is only child, and indicators for if one or both parents are missing. Peer controls is the reported percent of peers that the respondent expects to go to college.

-Marginal effects; Robust standard errors in parentheses;

-* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Completion

The completion model is also estimated by maximum likelihood via a probit model. Similar covariates to the first model are used and, for similar reasons as above, three models are estimated and analyzed. For the 1997 cohort, all students are used as none had completed any college at the time they were asked to report their expectations. For the 2001 cohort, a model is estimated that uses the entire cohort and controls for

¹⁴Whites were estimated to be less likely to attend college than Hispanics holding all other factors constant. This could be that after controlling for all of these factors there are more minority scholarships available that allow Hispanics to attend college at a higher rate. The Western United States also has large fraction of Hispanics that go and complete a two year college. This does not persist for the completion estimates.

¹⁵The appendix also includes estimates that used those from the 2001 sub-sample that were still enrolled in high school, along with a model that examines the number of years of college they completed.

enrollment.¹⁶ The last specification utilizes only the 700 students who answered the questions in both periods to examine how the predictive power of expectations changes for students over time. Again, Φ represents the cumulative distribution function of the standard normal distribution, t represents the cohort used in the estimation, and the other estimated coefficients have been suppressed for simplicity.

$$Prob(Complete\ 4\ Yr\ Degree_{i,t+n}) = \Phi(Expectations_{i,t}, X'_{i,t}\beta) \quad (3.2)$$

Columns 4 to 6 in table 3.6 above contain the estimates from the 1997 sub-sample. When the controls are added the coefficient diminishes considerably but remains positive and statistically significant. Also, the estimated impact on completion is smaller than for attendance but not as much as one might expect. It suggests that a 10 percentage point increase in a student's expectations is associated with a 2.25 percentage point increase in the probability that they go on to complete a four-year degree. This is only about 1.5 percentage points lower than for attendance and is likely a byproduct of using completing at least 1 year of college as the measure for attendance considering the drop-out rate is much higher for students in the first year.¹⁷ The model fit also increases when the expectations are included in the set of controls.

Table 3.7 below contains the estimates from two specifications that use the 2001 sample. The first, whose estimates are in columns 1-3, do not contain enrollment controls. I'll focus the discussion primarily on the second specification which contains these controls, whose estimates are in columns 4-6, as it appears to be the best fit. In both cases expectations are estimated to be statistically significant predictors of col-

¹⁶A Heckman selection model was attempted, but I was unable to identify a factor that influences attendance but not completion.

¹⁷While it is only a 1 percentage point difference, it is nearly 40% smaller so there is a non-negligible difference.

lege completion. The most comprehensive model suggest that a 10 percentage point increase in expectations results in an increase in the probability that the student completes college by 3.86 percentage points even when controlling for current enrollment status. The enrollment controls are each positive, statistically significant, and follow a logical sequence with those enrolled at a four-year college having a larger impact on future completion of a four-year degree than those enrolled at a two-year school. The same is true when comparing enrollment at a two-year college and high school. When expectations are added, the estimated relationship of these enrollment controls diminishes suggesting that some of the information included in the expectations variable is associated with current schooling status. Also, the estimated association of expectations and completion from the 2001 sample is considerably larger than the one obtained from the 1997 sample which indirectly suggests that as students age and progress through schooling the predictive power of their expectations increase.

Table 3.7: College Expectations and Completion with 2001 Sample - Probit Model

	Y = 1 if completed a 4 year college degree					
	(1)	(2)	(3)	(4)	(5)	(6)
College Expectations_01	0.00719*** (0.000287)		0.00506*** (0.000325)	0.00719*** (0.000287)		0.00386*** (0.000361)
Enrolled in 2yr College (d)					0.258*** (0.0510)	0.149*** (0.0440)
Enrolled in 4yr College (d)					0.607*** (0.0369)	0.421*** (0.0471)
Demographic Controls	No	Yes	Yes	No	Yes	Yes
Parent Education Controls	No	Yes	Yes	No	Yes	Yes
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in High School	No	Yes	Yes	No	Yes	Yes
Peer and Teacher Controls	No	Yes	Yes	No	Yes	Yes
Observations	1902	1815	1815	1902	1815	1815
Pseudo R^2	0.254	0.334	0.433	0.254	0.448	0.492

Notes: Demographic controls include male, white, and black. Parent education include indicators for if one or both of the respondents' parents have some college or a four-year degree. Location controls include rural, 1 of 4 census tracks, and MSA. Family controls include number of members in the household under 18, whether student is only child, and indicators for if one or both parents are missing. Peer controls is the reported percent of peers that the respondent expects to go to college.

-Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

-* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

To confirm the above hypothesis, I estimate the completion model again using the 700 students who answered the expectations questions in both 1997 and 2001. In 1997, respondents were between 15 and 17 and thus of high school age. When asked five years later they should have graduated and either proceeded into college or the work force. The estimates are included in table 3.8 below and confirm the hypothesis from above. When the respondents are in high school, a 10 percentage point increase in their expectations is associated with a 2.9 percentage point increase in the probability that they complete a four-year degree. Five years later when they report their expectations, the estimate increases to 3.8 percentage points. The last specification, whose estimates are included in column 7, include both the 1997 and

2001 expectations. Only the student's most recently reported expectations are found to be statistically significant. These results indicate that the predictive power of expectations increases as students age and graduate from high school.

Table 3.8: College Completion of Overlapping Sample in 1997 and 2001 - Probit Model

	Y = 1 if completed a 4 year college degree						
	1997 Sample			1997 Sample 5 Years Later			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
College Expectations_97	0.00686*** (0.000704)		0.00290*** (0.000754)				0.000296 (0.000599)
College Expectations_01				0.00700*** (0.000559)		0.00380*** (0.000569)	0.00378*** (0.000584)
Enrolled in 2yr College (d)					0.289*** (0.0715)	0.157* (0.0621)	0.157* (0.0624)
Enrolled in 4yr College (d)					0.645*** (0.0494)	0.391*** (0.0683)	0.389*** (0.0690)
Demographic Controls	No	Yes	Yes	No	Yes	Yes	Yes
Parent Education Controls	No	Yes	Yes	No	Yes	Yes	Yes
Location Controls	No	Yes	Yes	No	Yes	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes	Yes
Grades in High School	No	No	No	No	Yes	Yes	Yes
Peer and Teacher Controls	No	Yes	Yes	No	Yes	Yes	Yes
Observations	700	663	663	700	651	651	651
Pseudo R^2	0.135	0.355	0.370	0.299	0.551	0.595	0.596

Notes: Demographic controls include male, white, and black. Parent education include indicators for if one or both of the respondents' parents have some college or a four-year degree. Location controls include rural, 1 of 4 census tracks, and MSA. Family controls include number of members in the household under 18, whether student is only child, and indicators for if one or both parents are missing. Peer controls is the reported percent of peers that the respondent expects to go to college.

-Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1
 -* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The estimates above suggest that student expectations contain some form of private information that is not captured by the standard economic controls. These findings demonstrate that even at a young age students possess information regarding their potential for schooling success that econometricians have not yet considered. They also show that students gain relevant information regarding their true potential to complete a four-year college degree as they age. They use this information to form

more accurate predictions of their ability to complete a college degree. Below I test several other ways in which expectations might influence both college attendance and completion.

3.5 Sensitivity Analysis

This section examines other possible variations in which student expectations and post-secondary educational outcomes are related and what, if any, role parent expectations might play in student achievement. It is possible that the predictive power of expectations might vary across race and gender. I test this hypothesis by interacting expectations with gender and race. I then test a quadratic and cubic relationship. Following this I include indicators at the heaping points of 0, 50, and 100%. Last, I include parent expectations in the models and discuss the findings. Again, I use a probit model to examine each outcome and discuss the marginal effects evaluated at the means.

Expectations Interactions and Quadratic and Cubic Expectations

Previous behavioral literature found that men tend to be more overconfident in their abilities for future success, so it is plausible that women might be "better" predictors of future academic success. I test this by interacting expectations and gender and include it in the models. In each case I do not find evidence that women are better predictors. Also, the initial overconfidence by black males suggested possible differences across race. Race and expectations interactions were included in a separate specification, but were found to be insignificant. A specification which included both race and gender interactions was also tested and similar results were found.

It is plausible that the relationship between student expectations might be quadratic or cubic. Thus, I use both a quadratic and cubic specification to test this theory. The estimates from each are included in the table below. For attendance, columns 1 and

2 contain the quadratic and cubic specifications respectively. The completion estimates for the 1997 and 2001 cohorts are in columns 3-4 and 5-6 respectively. For each outcome, across all specifications, the estimates suggest that a quadratic and cubic relationship is inappropriate.¹⁸

Table 3.9: Quadratic and Cubic Expectations, Attendance, and Completion

	<u>Y = 1 if completed at least 1 year of college</u>		<u>Y = 1 if completed a 4 year college degree</u>			
	<u>1997 Sample</u>		<u>1997 Sample</u>		<u>2001 Sample</u>	
	(1)	(2)	(3)	(4)	(5)	(6)
College Expectations_97	0.00523*** (0.00138)	0.000898 (0.00341)	0.00212 (0.00132)	-0.000516 (0.00324)		
Expectations ²	-0.0000129 (0.0000115)	0.0000926 (0.0000781)	0.00000107 (0.0000104)	0.0000608 (0.0000682)		
Expectations ³		-0.000000655 (0.000000482)		-0.000000354 (0.000000400)		
College Expectations_01					0.00272 (0.00201)	-0.00261 (0.00514)
Expectations ²					0.0000126 (0.0000171)	0.000135 (0.000108)
Expectations ³						-0.000000726 (0.000000624)
Demographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Parent Education Controls	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes
Family Controls	Yes	Yes	Yes	Yes	Yes	Yes
Grades in 8th Grade	Yes	Yes	Yes	Yes	No	No
Grades in High School	No	No	No	No	Yes	Yes
Peer and Teacher Controls	Yes	Yes	Yes	Yes	Yes	Yes
College Enrollment Controls	No	No	No	No	Yes	Yes
Observations	3403	3403	3366	3366	1815	1815
Pseudo R ²	0.286	0.286	0.340	0.340	0.484	0.485

Notes: Demographic controls include male, white, and black. Parent education include indicators for if one or both of the respondents' parents have some college or a four-year degree. Location controls include rural, 1 of 4 census tracks, and MSA. Family controls include number of members in the household under 18, whether student is only child, and indicators for if one or both parents are missing. Peer controls is the reported percent of peers that the respondent expects to go to college.

-Marginal effects; Robust standard errors in parentheses

-* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

¹⁸The estimates from a linear probability model also suggest the findings above. The full model estimates for each specification can be found in the appendix.

Heaping or Confidence Indicators

I anticipate that those who reported a 0, 50, or 100% chance of obtaining a degree differed significantly from those who did not due to over or under-confidence, a lack of understanding of probability, or because they lacked focus or dedication when answering the survey question. Again, I estimate the attendance model using only the 1997 sample and the completion model using both. The estimates are presented in the table below. The only marginally significant coefficient from the heaping indicators comes from the model estimating degree completion that uses the 2001 sample. In this case, those who report a 100% chance of completing a four-year degree are estimated to be less likely to go on to complete it than those who do not. Specifically, those who report a 100% chance of obtaining a degree are 4.343 (.587-4.93) percentage points less likely to go on to complete the degree than those who report a 99% chance.

Table 3.10: Expectations Indicators, Attendance, and Completion

	Y = 1 if completed at least 1 year of college	Y = 1 if completed a 4 year college degree	
	1997 Sample	1997 Sample	2001 Sample
	(1)	(2)	(3)
College Expectations_97	0.00397*** (0.000570)	0.00278*** (0.000528)	
Expect=0_97 (d)	-0.0161 (0.0617)	0.109 (0.0803)	
Expect=50_97 (d)	-0.0430 (0.0309)	0.00576 (0.0301)	
Expect=100_97 (d)	-0.0462 (0.0288)	-0.0143 (0.0185)	
College Expectations_01			0.00587*** (0.000805)
Expect=0_01 (d)			0.309 (0.166)
Expect=50_01 (d)			0.0143 (0.0591)
Expect=100_01 (d)			-0.0493* (0.0223)
Demographic Controls	Yes	Yes	Yes
Parent Education Controls	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes
Family Controls	Yes	Yes	Yes
Grades in 8th Grade	Yes	Yes	No
Grades in High School	No	No	Yes
Peer and Teacher Controls	Yes	Yes	Yes
College Enrollment Controls	No	No	Yes
Observations	3403	3366	1815
Pseudo R^2	0.286	0.340	0.487

Notes: Demographic controls include male, white, and black. Parent education include indicators for if one or both of the respondents' parents have some college or a four-year degree. Location controls include rural, 1 of 4 census tracks, and MSA. Family controls include number of members in the household under 18, whether student is only child, and indicators for if one or both parents are missing. Peer controls is the reported percent of peers that the respondent expects to go to college.

-Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

-* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Parent Expectations

In the first round of interviews, both the selected youth and one of their parents each participated in a one hour long personal interview. During this interview they also asked the parent about the likelihood that their child would go on to complete a four-year college degree by the time they turned 30. It is possible that at such a young age, 15 to 17 when answering the questions, parents might possess more information regarding the probability that their child will go on to obtain a college degree. Since

the parents are only included in the initial survey, only those students who answer in both periods will be available from the 2001 sample. For this reason I estimate four models. The first two use the full 1997 sample to examine matriculation and completion. The second two limit the sample to those who answer in both periods and focus on college completion. When the limited sample is used, the first specification includes the students' expectations reported in 1997, while the second includes their expectations in 2001. This was done to examine how the roles of each change as the student ages.

The estimates from the models containing parent expectations are below. The addition of parent expectations results in a number of changes. For attendance, the estimated predictive power of the students' expectations decreases by .06 percentage points from the initial estimates, but the overall model fit increase.¹⁹ Also, the predictive power of parent expectations is nearly equivalent to the students. A ten percentage point increase in the student's expectations suggests an increase in the probability of completing at least one year of college by 3.1 percentage points. A similar increase in expectations for the parent results in a 2.76 percentage point increase. Examining completion, when the parent expectations are included using both the full and limited 1997 sample, the estimated predictive power of the parents' expectations dominates the expectations of the student. For example, utilizing the full sample estimates, a ten percentage point increase in expectations for the student increases the probability they complete a college degree by 1.82 percentage points, while a similar increase for the parent results in an increase of 2.08 percentage points. However, when the student is asked five years later, the parents' expectations have zero predictive power while the students' expectations remain positive and statistically significant.

¹⁹The .07 percentage point decrease is an 17% percent decrease.

Table 3.11: Parent Expectations and College Achievement

	Y = 1 if completed at least 1 year of college	Y = 1 if completed a 4 year college degree		
	Full 1997 Sample	Full 1997 Sample	1997 and 2001 Respondents	
	(1)	(2)	(3)	(4)
College Expectations_97	0.00310*** (0.000424)	0.00182*** (0.000379)	0.00215* (0.000898)	
College Expectations_01				0.00365*** (0.000528)
Parent Expectations_97	0.00276*** (0.000419)	0.00208*** (0.000360)	0.00306*** (0.000916)	0.000912 (0.000629)
Demographic Controls	Yes	Yes	Yes	Yes
Parent Education Controls	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes
Family Controls	Yes	Yes	Yes	Yes
Grades in 8th Grade	Yes	Yes	Yes	No
Grades in High School	No	No	No	Yes
Peer Controls	Yes	Yes	Yes	Yes
College Enrollment Controls	No	No	No	Yes
Observations	2903	2870	569	542
Pseudo R^2	0.304	0.353	0.400	0.601

Notes: Demographic controls include male, white, and black. Parent education include indicators for if one or both of the respondents' parents have some college or a four-year degree. Location controls include rural, 1 of 4 census tracks, and MSA. Family controls include number of members in the household under 18, whether student is only child, and indicators for if one or both parents are missing. Peer controls is the reported percent of peers that the respondent expects to go to college.

-Marginal effects; Robust standard errors in parentheses

-* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The findings suggest that parents also possess some valuable information regarding their child's schooling preferences and ability not captured by either the student's own expectations or the standard economic controls. The predictive power diminishes as the student ages which suggests that students are learning some of the private information that their parents have and are utilizing it when making predictions about their future success. Also, although student and parent expectations are highly correlated the addition of parent expectations does not drive the estimated predictive power of the students' expectations to zero.²⁰ This further indicates that students

²⁰Parent and student expectations are highly correlated. The correlation coefficient between the 1997 parent and student expectations is .57. It is .40 between parent expectations in 1997 and the student expectations in 2001.

have some form of private information that is orthogonal to the standard economic controls and the educational expectations of their parents when it comes to predicting their post-secondary success.

3.6 Conclusions

The recent increase in the cost of attending college has led to a number of discussions, including how to reduce these costs and whether a degree is still a worthwhile investment. However, the matriculation decision process and elements which contribute to both students' attendance and success are seldom discussed. I present new estimates of the factors that predict a student's decision to attend college and whether or not they complete a four-year degree that economists have not previously explored. I focus on the predictive power of students' self-reported probability of obtaining a four-year college degree. Later I consider how the predictive power of these expectations vary as the student ages and progresses through different levels of schooling. We anticipate that as students gain more information about the costs and benefits of the degree as well as their ability to complete it, their expectations should play a larger role in predicting their outcomes.

Manski (1993) outlined the importance of understanding and identifying how a student's expectations influence their decision to attend college or not. He notes that the estimated returns to education could be flawed when economists rely on the assumption of rational expectations instead of attempting to understand the process by which students use information to form expectations about the costs and benefits of a college degree. Dominitz and Manski (1996) and Manski (2004) both examine these issues, but are limited by small sample sizes and cross-sectional data. Jacob and Wilder (2011) are able to observe students over time and study the role of expectations in degree attainment, but do not examine whether these also influence the individual's decision to attend college. They are also limited by the nature of their data as their

measure of expectations is imprecise. This study improves upon these by utilizing data that follows students as they transition from high school into either college or the workforce and observes whether or not they obtained a four-year college degree by the time they turn thirty.

A unique set of questions from the National Longitudinal Study of 1997 (NLSY97) allows for an examination of how a student's subjective probability of completing a college degree changes as they progress through schooling and what role it plays in their outcomes. The survey asks a subset of students to report their probability of obtaining a four-year college degree by the time they turn 30, once in high school and again when some have entered either college or the workforce. To date no other studies have used these data to examine the role of expectations in educational outcomes. The probability scale is an improvement over other studies as most use data where students respond on a Likert scale regarding their educational expectations, or are simply asked the highest level of education they expect to complete without any notion of how confident they are in realizing that outcome.

The results indicate that, even after controlling for the factors previously identified in the literature, a student's self-reported probability of completing a four-year college degree plays a role in predicting whether they attempt to complete a college degree and if they are successful in obtaining a four-year degree. As students age and progress through more schooling, their expectations play a larger role in predicting completion. Also, parent expectations are estimated to be influential when students are young, but the impact diminishes as the student ages and gathers information related to the costs and benefits of a college degree. The estimates signal that students' expectations contain a form of private information that is helpful in determining post-secondary schooling attendance and success that isn't being captured by the previously identified controls, and suggests that youth and econometricians may possess and utilize different data when analyzing schooling decisions. While the

combination of interpretable subjective data on expectations and student choice data could help solve the identification problem discussed by Manski (1993), further analysis is needed to discover how students are forming and updating these expectations.

Understanding the process by which students both form and update their expectations is essential to understanding the underlying role that these expectations play in the attainment process. The use of interpretable subjective data on expectations, combined with choice data, can help solve the identification problem in the returns to education literature by relaxing the rational expectations assumption. To accomplish this, a study that examines how students form and update their expectations about whether or not to attend college is necessary and is included below. Lastly, if the information that makes students more accurate predictors of their future success can be identified, it is possible we could also increase the matriculation and graduation rates at post-secondary schools. This is analyzed in the third essay.

Chapter 4 The Formation and Updating of Post-secondary Educational Expectations as Student's Transition from High School

4.1 Introduction

Obtaining reliable estimates of the returns to education requires that researchers successfully deal with the issues of selection. A popular approach is to model school choice using observed outcome data, assume that students have rational expectations, and that they condition their beliefs on similar variables and process information in a similar way (Manski, 1993). However, students make schooling choices under uncertainty - uncertainty about personal tastes, individual abilities, and realizations of choice-related outcomes (Zafar, 2011). While the decision process has been examined theoretically (Manski, 1989; Altonji, 1993; Hilmer, 1998; Malumud, 2005) and empirically (Bamberger, 1987; Arcidiacono, 2004; Strange, 2012; Hunter, 2017a), most rely implicitly or explicitly on the unrealistic assumptions noted above regarding the expectations of students. If these assumptions are violated then a number of issues arise. First, if we are unaware of how youth perceive the returns to schooling, then it is impossible to infer the decision process from their observed schooling choices. Second, when not knowing a youth's decision process, inferring the objective returns to schooling from data on realized outcomes is implausible. These issues make an examination of how students form and update their expectations about their future education essential.

This paper has two goals, the first is to identify whether or not students incorporate the factors associated with college success into their expectation that they will complete college. The second is to test whether or not students update their educational expectations in a Bayesian manner. A unique question regarding the expectation that participants will go on to complete a four-year college degree, along

with detailed responses from both the participants and their parents, from the 1997 National Longitudinal Survey of Youth (NLSY97) are used to investigate both topics. It is a nationally representative data set following individuals over time. Unlike previous studies, I observe students first while they are in high school and then follow them on an annual basis. This allows me to observe how their expectations change as they age and advance out of high school. I also am able to observe whether they fulfill their expectations. Their outcomes also are used to estimate whether or not they are Bayesian.

I begin by estimating which factors students appear to be utilizing when forming their educational expectations. I do this first for a subset of individuals in the NLSY97 that were between 15 and 17 years old when asked, and then later for another subset who were between 17 and 22 years old. I then compare the information that students appear to be using to form their educational expectations to factors that have been previously shown to impact college completion. I find that while high-school aged students are taking into account several of the relevant factors when forming their expectations, they also appear to be neglecting the impact that income and ability have on their likelihood of completing college. Overall, students are extremely overconfident in their ability to complete a four-year college degree by the time they are 30 possibly because they are not incorporating these factors; they could also be classified as overly aspirational based upon their observed characteristics.

To test whether or not students are Bayesian, I use both samples mentioned above as well as 710 students who reported their expectations in the first round (1997) and then again 5 years later in the fifth round (2001). I develop a Bayesian model and discuss three predictions for how students should respond to the acquisition of new information. I identify four sources of information that became available to students between 1997 and 2001 and test whether or not students responded to the new information in a way that the model would predict. I show that while students

do appear to be utilizing the new information when updating their expectations, they do not appear to be true Bayesian updaters. I hypothesize two possibilities as to why this is not the case.

4.2 Formation of College Completion Expectations

Summary statistics for how the three subsets responded to the college expectations question are contained in table 4.1. On average, respondents in each sub-sample are on average overconfident, or more aspirational, in their ability to complete college. The average expectation for the 1997 sub-sample was 72.79% but only 25% of them complete a four-year degree by 2013. The older, 2001 sub-sample had an average expectation of 67.17% and only 27% of them completed college by 2013.¹ The 2001 sub-sample also had more variation than the 1997 sub-sample. The lower expectations for the older cohort suggests that individuals are gathering relevant information about their true ability to complete college as they age and are incorporating this information into their expectations by updating downward. This hypothesis is supported by examining the 710 who reported in both 1997 and then 5 years later in 2001. On average, the 710 respondents decreased their expectations by 9.35 percentage points and the variation increased substantially. Both the decrease in expectations and increase in variation is due to students becoming less certain they will be able to complete a degree. More students are updating downward over the period and a much larger portion report a 0% chance of completing a degree while the proportion who report a 100% chance is relatively unchanged. A more detailed analysis of how these expectations change is included in a later section. Although students are updating downward, on average they still lack some relevant information as they remain extremely overconfident in their ability complete a four-year college degree.

¹This is nearly identical to the national statistics of college completion for students during this period.

Table 4.1: Expectations of Obtaining a Four-year College Degree

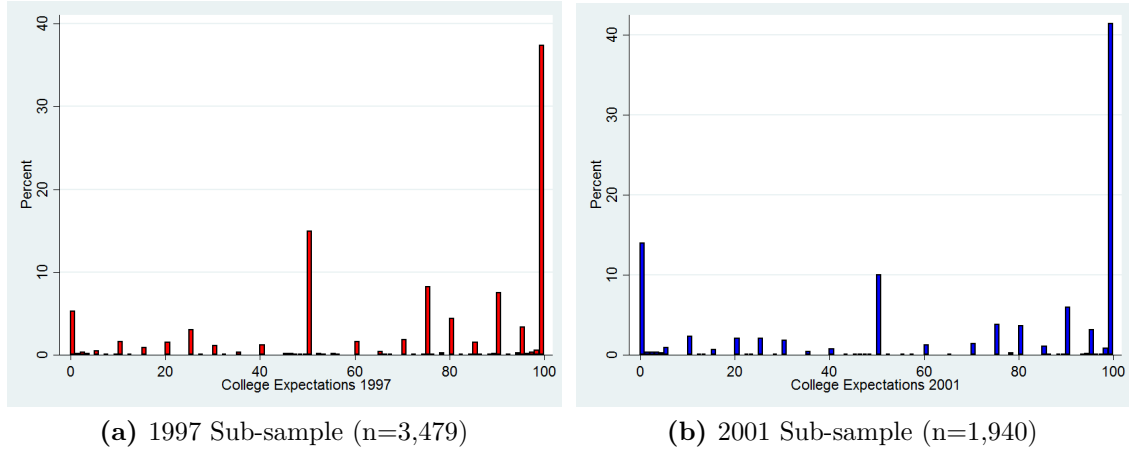
	N	Mean	Standard Deviation	Minimum	Maximum
1997 Sub-sample	3,479	72.79%	31.43	0	100
2001 Sub sample	1,940	67.17%	38.48	0	100
Overlapping Sample					
In 1997	710	74.29%	30.81	0	100
In 2001	710	64.94%	39.88	0	100
Δ Expectations	710	-9.35	39.49	-100	100

Notes: The standard deviation is percentage points and not percent.

The distribution of the reported educational expectations for both the 1997 and 2001 sub-samples are also included below. These show that there is considerable variation in the expectation that individuals have that they will complete a four-year college degree by the time they are 30. The respondents used the entire range from 0 to 100, but generally rounded their answers to the nearest 5 especially for responses not at the extremes.² There is also considerable heaping at the certainty points of 0% and 100% and the uncertainty point of 50% in both sub-samples. Comparing the two, a larger percentage of respondents reported at the certainty points and less at the uncertainty point in the 2001 sub-sample. This also suggests that older individuals have more information about their true likelihood of completing college and are using this to form more certain beliefs about their ability to complete a degree.

²This is similar to findings from Zafar (2013)'s study of Northwestern students' expectations of future academic outcomes.

Figure 4.1: Distribution of College Completion Expectations in 1997 and 2001



This pattern is also present in the responses of the 710 who reported their expectations in both 1997 and 2001. Table 4.2 presents how these individuals responded in each period. Examining the distributions for both years it appears the majority of changes are for students who are learning that they do not have the ability to complete a college degree and are updating accordingly. We observe a small increase in the number of students who reported 100%, a large increase in the number who reported 0%, and a small decline in those who reported a 50% chance. In 1997, 265, or approximately 37%, of the respondents were certain they would complete a college degree, and in 2001, 284, or 40%, were 100 percent certain. Only 25, or 3%, reported a 0 percent chance of completing a college degree in 1997, but five years later, 120, or almost 17%, of the respondents reported a 0 percent chance. Overall, 222 of the students did not update their expectations, 209 increased their expectations, and 279 decreased their expectations.³

³A table that outlines the direction and magnitude of how respondents updated their expectations is included in the appendix.

Table 4.2: Frequency Distribution of Expectations in 1997 and 2001

1997 Expectations				2001 Expectations		
	Freq	%	Cum. %	Freq	%	Cum. %
0	25	3.52	3.52	120	16.9	16.9
1	1	0.14	3.66	2	0.28	17.18
2	4	0.56	4.23	1	0.14	17.32
3	1	0.14	4.37	2	0.28	17.61
4	1	0.14	17.75
5	4	0.56	4.93	8	1.13	18.87
10	17	2.39	7.32	16	2.25	21.13
15	5	0.7	8.03	3	0.42	21.55
20	9	1.27	9.3	9	1.27	22.82
22	1	0.14	22.96
23	1	0.14	23.1
25	26	3.66	12.96	17	2.39	25.49
27	1	0.14	13.1
30	8	1.13	14.23	15	2.11	27.61
32	1	0.14	14.37
35	3	0.42	14.79	2	0.28	27.89
40	7	0.99	15.77	5	0.7	28.59
43	1	0.14	28.73
46	1	0.14	15.92	1	0.14	28.87
50	91	12.82	28.73	80	11.27	40.14
55	1	0.14	40.28
60	16	2.25	30.99	8	1.13	41.41
65	3	0.42	31.41	2	0.28	41.69
70	16	2.25	33.66	4	0.56	42.25
72	1	0.14	33.8
75	61	8.59	42.39	20	2.82	45.07
76	1	0.14	42.54
78	2	0.28	42.82	1	0.14	45.21
80	27	3.8	46.62	22	3.1	48.31
85	12	1.69	48.31	8	1.13	49.44
89	1	0.14	48.45	1	0.14	49.58
90	49	6.9	55.35	36	5.07	54.65
92	1	0.14	55.49
94	1	0.14	54.79
95	25	3.52	59.01	14	1.97	56.76
96	1	0.14	59.15	1	0.14	56.9
97	4	0.56	59.72
98	5	0.7	60.42	3	0.42	57.32
99	16	2.25	62.68	19	2.68	60
100	265	37.32	100	284	40	100

Notes: Included above are only the 710 who answered the educational expectations question in both 1997 and 2001. In 1997, students were between 15 and 17 when reporting the percent chance they would complete a four-year college degree. In 2001 they were between 19 and 22.

Prior to examining what information students might be using to update their expectations, I use OLS to examine how these expectations vary conditional on different demographics, household structure and characteristics, location, schooling success and ability, and the child's parental expectations of college success.⁴ The estimates

⁴Davies and Kandel (1981) and Teachman and Paasch (1998) support the inclusion of the parent's expectations as education and income are found to only account for 1/3 of the family contribution to college aspirations.

from several specifications utilizing the 1997 sub-sample and the overlapping sample are included in table 4.3. Only two factors were found to be significantly associated with educational expectations across all specifications for those between 15 and 17 years old. Those who earned Bs or better while in 8th grade and those who had at least one parent in the household with a college degree reported higher expectations; parent expectations were also estimated to be statistically significant in each case.⁵ This suggests that students are basing their expectations of future academic success on their own previous academic success, whether or not they have at least one parent with a college degree, and what their parent's expectations are of them in terms of completing college. It also indicates that students possibly lack information, have imperfect information, or simply are not incorporating other factors associated with college success such as family income and ability into their expectations. To test this I included the same covariates in a probit model where the dependent variable was 1 if the student completed college by 2013 and 0 otherwise. The estimates from this are included in columns 5 and 6 of table 4.3 and are discussed in more detail later.⁶

⁵The NLSY97 did not ask high school grades until 1998 so they are not included in the information set of the students reporting their expectations in 1997.

⁶This is a similar model to essay one but excludes the students' expectations.

Table 4.3: 1997 Expectation Formation and College Completion

	Y = Expectation of completing a college degree in 1997				Probit for Completion of College Degree	
	(1)	(2)	(3)	(4)	(5)	(6)
Parent's College Expectations		0.416*** (0.0207)		0.420*** (0.0479)	0.0114*** (0.00135)	0.0152*** (0.00318)
Age in 1997	-3.573*** (0.845)	-1.265 (0.834)	-0.373 (1.903)	0.931 (1.861)	-0.0799 (0.0541)	0.115 (0.121)
Male	-4.645*** (0.963)	-4.944*** (0.946)	-4.356* (2.163)	-3.801 (2.127)	-0.118 (0.0617)	-0.150 (0.144)
White	-5.266** (1.601)	-1.544 (1.555)	-6.325 (3.622)	-5.437 (3.525)	0.0210 (0.110)	0.143 (0.233)
Black	6.617*** (1.550)	4.458** (1.570)	6.176 (3.374)	2.119 (3.408)	0.337** (0.110)	0.442 (0.249)
Income Quantile	0.914 (0.628)	-0.169 (0.589)	1.087 (1.395)	-0.00571 (1.249)	0.114** (0.0395)	0.0295 (0.0901)
As in 8th	17.65*** (1.858)	10.07*** (1.888)	20.83*** (3.940)	12.67*** (3.719)	0.923*** (0.130)	1.254*** (0.292)
As & Bs in 8th	13.04*** (1.801)	7.263*** (1.813)	19.20*** (3.933)	11.84** (3.780)	0.490*** (0.118)	0.490 (0.264)
Bs in 8th	10.80*** (1.928)	7.653*** (1.944)	15.92*** (4.314)	9.012* (4.319)	0.358** (0.125)	0.534 (0.282)
Bs & Cs in 8th	3.573* (1.802)	2.434 (1.788)	7.683 (3.990)	6.046 (3.754)	0.0875 (0.121)	0.0610 (0.264)
Cs & Ds in 8th	-5.696* (2.446)	-1.445 (2.463)	-3.289 (5.635)	-0.253 (5.205)	-0.172 (0.195)	-0.244 (0.438)
No HS Degree for both Parents	-3.544 (2.024)	-4.247* (2.066)	-1.458 (4.677)	-2.454 (4.439)	-0.0393 (0.144)	0.154 (0.308)
One Parent has Some College	5.111*** (1.170)	1.961 (1.158)	5.276* (2.612)	4.315 (2.680)	0.0375 (0.0722)	0.182 (0.162)
Both Parents have Some College	3.774* (1.915)	4.580** (1.775)	5.727 (3.981)	0.458 (4.095)	0.233 (0.124)	0.0674 (0.282)
One Parent has a College Degree	10.64*** (1.246)	7.322*** (1.187)	10.89*** (2.949)	6.819* (2.825)	0.465*** (0.0799)	0.827*** (0.193)
Both Parents have a College Degree	2.715 (1.480)	0.950 (1.389)	4.267 (3.226)	3.180 (3.096)	0.257* (0.117)	-0.114 (0.276)
ASVAB Percentile	0.119*** (0.0227)	0.0346 (0.0224)	0.102 (0.0531)	0.0174 (0.0545)	0.0104*** (0.00143)	0.00779* (0.00338)
Highest Grade Completed 1997	2.376*** (0.650)	0.843 (0.651)	0.721 (1.504)	-0.983 (1.443)	0.100* (0.0436)	0.0969 (0.0962)
Enrolled in HS in 1997	24.16*** (2.678)	14.23*** (2.655)	25.16*** (6.787)	9.801 (5.990)	0.257 (0.269)	0.514 (0.487)
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes
Family Structure Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3386	2905	689	601	2905	601
Adjusted R^2	0.255	0.389	0.248	0.357		

Notes: Location controls include: whether or not the family lived in an MSA and if so if it were in the central city or not; whether they lived in an urban or rural area; and which of the four large standard census regions they lived. Controls for household structure include: whether the mother, father, or both were absent from the household; if the student were an only child; and the number of other siblings under 18 in the household in 1997.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

—Robust standard errors in parentheses

Examining the expectation formation estimates, the largest changes occur when the parents' expectations of their child's college completion are included in the model. Columns 1 and 2 contain the estimates using the 1997 sub-sample first without and then with the parents' expectations. Prior to the inclusion of the parents' expectations, several characteristics are statistically significant. Those who are older, male, or white all had lower expectations of completing college. Those who are black, earn

better grades in 8th grade, are enrolled in high school, have at least one parent with some college experience or more, or score higher on the ASVAB had higher expectations on average. Excluding age and race, in relative terms expectations appear correlated with the completion rates of students both in this data and from national statistics of the period.⁷ Those who achieve higher marks in school and who have parents with more education go on to complete college at a higher rate. Also, men graduate at a lower rate than women so the differences in expectations based on these factors suggests that students understand this; excluding gender these factors are confirmed as increasing the likelihood that the student completes college by the probit model.

When the parent's expectations are included they are positive, statistically significant, increase the explanatory power of the model, and alter the significance and magnitude of several of the other factors. This indicates that the students are forming their expectations in large part based upon what their parents' aspirations are of them and that their parents might understand or have more access to information that predicts college success. The findings suggest that children want to fulfill their parent's expectations of them or it is possible that they assume their parents have more complete information about what is needed for them to complete a college degree and are thus incorporating it into their own expectations.⁸ The age of the student, whether or not they are white, and their ASVAB percentile no longer are statically significant with the inclusion of the parents' expectations which indicates that these are potentially capturing some of this information indirectly. It might also

⁷Blacks having higher expectations than Hispanics is somewhat counter-intuitive as they complete college at a similar rate. This is driven primarily by the high expectations reported by black males while in high school. Hispanics having higher expectations than whites is also counter-intuitive when comparing the completion rates of each group. This is also counter-intuitive to findings from Stewart et al. (2007).

⁸The interviews of the parent and child were separate so neither should have known what the other reported.

signal that the students are using their parents' expectations as a source of information about their ability. The individual's ASVAB score is not revealed to them until several weeks/months after the initial survey. Later I utilize the ASVAB score as a new source of information available to students when examining how they are updating their expectations. The impact of being enrolled in high school also diminishes but remains statistically significant. These changes also indicate that parent expectations are highly correlated with a number of the relevant factors and suggests that parents are utilizing this information when forming expectations of their child's future success, further analysis is needed to disentangle the two possibilities.

Columns 3 and 4 contain the estimates from the overlapping sub-sample first without and then with the parents' expectations. The coefficient estimates in both are similar and indicate that in addition to parent expectations, good grades in 8th grade and having at least one parent with a college degree are positively associated with a student's educational expectations. Similar to the model that utilizes the 1997 sub-sample, when the parents' expectations are included they are statistically significant, the explanatory power of the model increases, and the estimated impact of other factors diminishes. This reinforces the notion that these parent expectations are highly correlated with a number of the other factors and suggests that parents are utilizing this information when forming expectations about child's ability to complete a four-year college degree by the time they are 30. Future research will examine what factors parents utilize when they form their expectations of college success for their child.

Both previous research and the model of college completion estimated above show that students who have more previous academic success, a higher family income, parents with more education and who have higher aspirations for them, and who have higher ability measures are more likely to go on to complete a four-year college

degree.⁹ The estimates suggest that high school students, although overconfident in their ability to complete a college degree, are incorporating some relevant information when forming their expectations about future college success.¹⁰ However, they also appear to be either neglecting or failing to incorporate other useful predictors of college success, in particular family income and ability. It is possible that students are unaware of how income is related to college success, and might be uncertain about their true ability and how this translates to their likelihood of completing college. There are also noticeable difference in expectations based on gender and race that are concerning as they do not appear directly correlated with college completion after the standard economic controls are included. In the next section I examine how students' expectations change as they age and transition out of high school and test whether students learn about their ability and its relationship with college success.

4.3 The Change in Students' Expectations from 1997 to 2001

Above we have shown that high school aged students are incorporating some of the primary predictors of college success into their expectations about being able to complete college, but there are also several factors they appear to be neglecting. In this section we explore how students update their expectations as they age and whether or not they ultimately incorporate the information they appear to be neglecting initially. I first focus on how expectations change for the 710 who answered in both 1997 and 2001. A Bayesian model of student updating is then presented and discussed. I present four new sources of information that became available to the

⁹Location controls include: whether or not the family lived in an MSA and if so if it were in the central city or not; whether they lived in an urban or rural area; and which of the four large standard census regions they lived. Controls for household structure include: whether the mother, father, or both were absent from the household; if the student were an only child; and the number of other siblings under 18 in the household in 1997.

¹⁰Hanson (1994) also discusses some of these differences in college aspirations and findings are similar.

individuals between 1997 and 2001 and test whether or not students appear to be Bayesian updaters.

Figure 4.2 presents the distribution of responses from the 710 who reported their expectations in both periods. Students used the entire scale and there is considerable heterogeneity in beliefs. When comparing the two, we see the majority of movement is toward the lower bound of zero. In 1997 only 3.5% of students reported a zero percent chance of completing a college degree, but in 2001 17% reported a 0 percent chance. Similarly, in 1997 only 29% of students reported less than or equal to a 50 percent chance while in 2001 nearly 40% reported in this range. The same is not true for the upper portion of the distribution. Approximately 37% of students reported a 100 percent chance of completing a college degree in 1997, while 40% reported a 100 percent chance in 2001. This pattern is consistent for those reporting greater than or equal to 75 percent chance.

Figure 4.2: Distribution of College Completion Expectations for the Overlapping Sub-sample

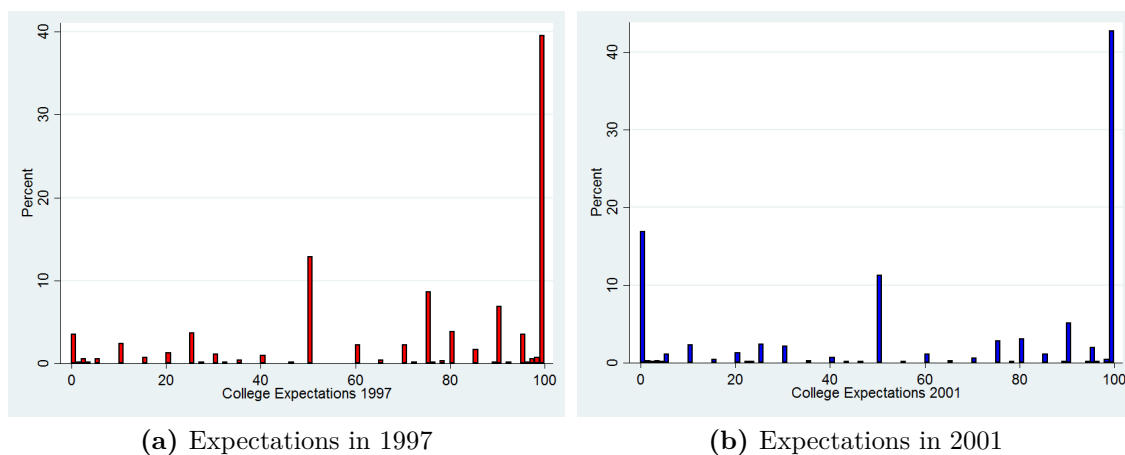


Table 4.4 presents both the direction and magnitude of how the students' expectations changed from 1997 to 2001. Overall, 209 of the 710 respondents updated positively, 279 updated negatively, and 222 reported the same expectations in both periods. For those who made revisions, the majority updated between 1 and 25 per-

centage points and the second largest category were those who updated between 26 and 50 percentage points. However, students did update across the entire spectrum with some students initially reporting a 0% chance of obtaining a college degree in 1997 and then a 100% chance in 2001 and vice versa. This indicates that students, or at least those who report at the two certainty points, might not be updating in a Bayesian fashion as the model predicts that those with priors of 0% or 100% should not update their expectations regardless of what information they receive. The next section presents a Bayesian model for students and tests whether or not they are Bayesian updaters.

Table 4.4: Direction and Magnitude of Updated Expectations from 1997 to 2001

Increase in Chance Have a 4 Year Degree by 30	Number of Individuals	Mean 1997 Expectations	Mean 2001 Expectations	Actual Completion
1 - 25 %	124	76.01%	89.31%	47.58%
26 - 50 %	59	44.49%	86.17%	22.03%
51 - 75 %	13	26.07%	93.08%	7.69%
76 - 90 %	7	8.81%	94.29%	0%
91 - 99 %	1	3%	100%	0%
100 %	5	0%	100%	20%
Total	209	59.56%	89.13%	35.41%
Decrease in Chance Have a 4 Year Degree by 30	Number of Individuals	Mean 1997 Expectations	Mean 2001 Expectations	Actual Completion
1 - 25 %	109	65.02%	51.11%	16.51%
26 - 50 %	74	73.62%	28.08%	0%
51 - 75 %	43	81.93%	12.72%	2.33%
76 - 90 %	20	91.35%	6.75%	0%
91 - 99 %	9	98%	1.89%	0%
100 %	24	100%	0%	4.17%
Total	279	75.87%	29.92%	7.17%
No Change in Chance Have a 4 Year Degree by 30	Number of Individuals	Mean 1997 Expectations	Mean 2001 Expectations	Actual Completion
Total	222	86.17%	86.17%	47.75%

Notes: Included above are only the 710 who answered the educational expectations question in both 1997 and 2001. In 1997, students were between 15 and 17 when reporting the percent chance they will complete a four-year college degree. In 2001 they were between 19 and 22.

4.4 Are Students Bayesian Updaters?

The remainder of the paper tests whether or not students are Bayesian updaters. I first outline a Bayesian model of student expectations and the necessary assumptions. I then use the model to make predictions of how students should update if they are Bayesian. I identify four potential sources of new information that became available to the student between 1997 and 2001 and hypothesize the impact they should have in the updating process. Last, I test the predictions using the 710 students who answered the expectations question in both 1997 and 2001.

Bayesian Model

$$Prob(S_{c,2001}|M_{2001-1997}) = \frac{Prob(M_{2001-1997}|S_c)}{Prob(M_{2001-1997})} \times Prob(S_{c,1997})$$

In the model above S_c represents the state where the individual completes a four-year college degree. $M_{2001-1997}$ represents the messages or new information that they receive between 1997 and 2001. The posterior expectation, $Prob(S_{c,2001}|M_{2001-1997})$, is the conditional (posterior) probability of completing college by 30, given they receive message M. In the data this is the reported expectation of completing college in 2001. The prior expectation, $Prob(S_{c,1997})$, is the unconditional (prior) probability of completing college by 30, and is the reported expectation of completing college in 1997. The components that are estimated from the data are those associated with how the individual process new information and the magnitude of the impact that it should have. The first, or $Prob(M_{2001-1997}|S_c)$, is the conditional probability of receiving message M, given the true state is that the individual completes college by 30. The second, $Prob(M_{2001-1997})$, is the unconditional probability of receiving message M. I discuss how both of these components are estimated from the data following a discussion of the necessary assumptions and the sources of new information.

There are two primary assumptions associated with this model. The first is that individuals are assumed to use all of the information available to them when forming their expectations and therefore revisions are a result of the acquisition of new information only. The second is that any information available to the individual in the past is also available to them at any point in the future.¹¹ Estimates from the model with expectations as the dependent variable indirectly support the first assumption as they indicate that when students are forming their expectations, although imperfectly, they are incorporating several factors related to college success. Since the student or parent reported the information in the survey, they must have had access to the information when forming their expectations. I test the second assumption by regressing the expectations in 2001 on the same covariates as in 1997. If this assumption holds, the estimated impact of each covariate and the percent of variation explained by the model should be similar. I estimate using both the 2001 sub-sample and the 710 who answered in both periods and compare the results.¹²

Columns 1 and 2 in table 4.5 contain the estimates from regressing the 1997 expectations on the same covariates as before, excluding the parents' expectations.¹³ Columns 3 and 4 contain the estimates from regressing the 2001 expectations on the same set of covariates. Column 1 includes the estimates from the full 1997 sub-sample, column 3 includes estimates from a model that utilized the entire 2001 sub-sample, and columns 2 and 4 are the estimates from the 710 who answered in both periods. I indirectly test assumption 2 by comparing the estimates from columns 1 and 3. If information was available and being utilized in 1997 then it also should be utilized

¹¹Zafar (2011, 2013) also use these two primary assumptions when testing whether or not students are Bayesian.

¹²The full 1997 and 2001 sub-samples can be loosely compared, but because of the differences in age and enrollment when surveyed it is best to compare the estimates for the 710 who answered in both periods.

¹³These are excluded as it limits the sample size. Only a portion of those who were asked their expectations in 2001 also had parents who answered the expectations questions.

in the formation of expectations in 2001. Although the samples differ in terms of age and enrollment, the estimates appear similar for the majority of the covariates. To further test the assumption that information in the past is also available in the future we compare columns 2 and 4 which uses the same set of students. A chi-square test of equality for each of the coefficients indicate that the majority are statistically similar.¹⁴ This indicates that the information available and utilized by the respondents in 1997 was also available and utilized in 2001. Regardless of specification or sub-sample used, the percentage of variation in the students' expectations that is explained by the model is between 24% and 25% which supports assumption 2, that any information available to the individual in the past is also available to them at any point in the future.

¹⁴The indicator for sample type, if the respondent's household lacked both parents, and ASVAB percentile were the only three that were statistically different at the 5% level.

Table 4.5: 1997 and 2001 Expectation Formation

	Y = College Expectations in 1997		Y = College Expectations in 2001	
	(1)	(2)	(3)	(4)
Age in 1997	-3.573*** (0.845)	-0.373 (1.903)	-4.320*** (1.157)	-4.178 (2.308)
Male	-4.645*** (0.963)	-4.356* (2.163)	-4.775** (1.584)	-4.207 (2.785)
White	-5.266** (1.601)	-6.325 (3.622)	-5.570* (2.673)	-6.182 (5.002)
Black	6.617*** (1.550)	6.176 (3.374)	5.229* (2.626)	7.042 (4.494)
Income Quantile	0.914 (0.628)	1.087 (1.395)	2.423* (1.040)	3.710 (1.907)
As in 8th	17.65*** (1.858)	20.83*** (3.940)	16.62*** (3.039)	19.36*** (5.267)
As & Bs in 8th	13.04*** (1.801)	19.20*** (3.933)	11.82*** (2.859)	14.35** (5.018)
Bs in 8th	10.80*** (1.928)	15.92*** (4.314)	10.27*** (3.056)	14.19** (5.252)
Bs & Cs in 8th	3.573* (1.802)	7.683 (3.990)	1.897 (2.885)	6.893 (4.897)
Cs & Ds in 8th	-5.696* (2.446)	-3.289 (5.635)	-8.348* (3.650)	-2.043 (6.607)
Ds in 8th	-5.797 (3.607)	0.277 (6.228)	-10.26 (5.430)	0.832 (8.685)
Fs in 8th	-8.323 (5.815)	-28.43** (10.15)	-12.45 (7.414)	-9.358 (14.57)
No HS Degree for both Parents	-3.544 (2.024)	-1.458 (4.677)	-4.372 (3.155)	-0.683 (5.529)
One Parent has Some College	5.111*** (1.170)	5.276* (2.612)	4.592* (2.068)	1.873 (3.582)
Both Parents have Some College	3.774* (1.915)	5.727 (3.981)	10.27** (3.199)	9.222 (5.745)
One Parent has a College Degree	10.64*** (1.246)	10.89*** (2.949)	13.92*** (2.194)	16.28*** (3.889)
Both Parents have a College Degree	2.715 (1.480)	4.267 (3.226)	4.408 (2.606)	0.614 (4.859)
ASVAB Percentile	0.119*** (0.0227)	0.102 (0.0531)	0.246*** (0.0409)	0.326*** (0.0729)
Highest Grade Completed 1997	2.376*** (0.650)	0.721 (1.504)	2.896** (1.106)	4.114* (1.979)
Enrolled in HS in 1997	24.16*** (2.678)	25.16*** (6.787)	11.16 (6.581)	10.04 (8.311)
Location Controls	Yes	Yes	Yes	Yes
Family Structure Controls	Yes	Yes	Yes	Yes
Observations	3386	689	1893	689
Adjusted R^2	0.255	0.248	0.248	0.245

Notes: Location controls include: whether or not the family lived in an MSA and if so if it were in the central city or not; whether they lived in an urban or rural area; and which of the four large standard census regions they lived. Controls for household structure include: whether the mother, father, or both were absent from the household; if the student were an only child; and the number of other siblings under 18 in the household in 1997.

-* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

-Robust standard errors in parentheses

Comparing columns 1 and 3 that utilize each sub-sample separately and columns 2 and 4 where only those who answered in both period, the biggest discrepancies are associated with the ASVAB percentile and the enrollment variables. The ASVAB percentile is estimated to have a larger impact in the formation of the 2001 expecta-

tions.¹⁵ The student did not receive these scores until after the first year interview, so this could potentially be a new source of information that students use to update their expectations.¹⁶ For enrollment, the majority of the students in the 2001 sub-sample and all of the 710 who answered in both periods should have graduated high school by 2001. Therefore, whether or not that had already completed high school and their current enrollment status at a post-secondary institution would dominate their previous high school enrollment as a source of information about the possibility that they could complete a college degree. Both the ASVAB score and college enrollment status in 2001 are tested as potential new sources of information that students use to update their expectations. Two other potential sources of information that became available to the individual between 1997 and 2001 include whether or not they took a college entrance examination, ACT or SAT, and what their reported high school grades were. These sources of information are discussed in more detail below.¹⁷

Potential Sources of Information:

1. Individual receives their ASVAB test scores¹⁸
2. Individual takes a college entry examination (ACT or SAT)¹⁹
3. After finishing High School, the individual can join the workforce, enroll in a two-year college, or enroll in a four-year college²⁰
4. Individual receives their High School grades which can differ from the grades

¹⁵The chi-square test of coefficients across the models that used the overlapping sub-sample indicate that the two are statistically different.

¹⁶They also received a breakdown of their percentile score for several different sub-tests and what each category related to in terms of future work possibilities.

¹⁷Detailed summary statistics for each measure can be found in the appendix.

¹⁸The ASVAB scores for some students were imputed to maximize sample size. These are used only for robustness checks.

¹⁹Taking either of these exams was used over the score in order to maximize sample size. The student's reported scores were not used as it also limited the sample size severely.

²⁰Only 11 students obtained an AA degree prior to 2001 so completion of this degree was not used as a source of information.

they received in 8th grade²¹

There are obvious issues of selection for those who take a college entry examination and who ultimately enrolls in these institutions, therefore they are only used in the initial descriptive analysis. When estimating the informational portion of the Bayesian model I focus primarily on the ASVAB score as there is limited variation in the number of students whose grades change from 8th grade to high school. The initial estimates also suggest that students are not using these changes as a primary source of information.

Predictions from a Bayesian Model

If students are updating in a Bayesian fashion then we should observe the following. First, the value of the new information should depend directly on the initial confidence or aspirations of the student. New information should be the most valuable to the most uncertain students and the least valuable to the most certain. Per the model, for those who believe with certainty that they either will or will not complete a college degree, any new information should be uninformative and their expectation should not change.²² To test this I examine whether or not those who reported either 0% or 100% in 1997 update their expectation in 2001. Overall, 25 students reported a 0% chance of completing college in 1997. Of those 25, only 14 also reported a 0% chance in 2001 and the average expectation of the 25 respondents increased by approximately 30 percentage points.²³ For the 265 students who initially reported a

²¹Although a large portion of these students were enrolled in high school in 1997, they were not asked to report their high school grades until 1998.

²²This is a basic tenant of a Bayesian model. For those who reported a 100% chance, the denominator in the impact and processing portion of the model simply reduces to the conditional probability of receiving message M, given the individual complete college by 30 and the result is simply one times the prior. If the prior were 0%, then anything times zero is zero and thus the posteriors should also equal zero.

²³The 11 respondents who increased their expectations did so on average by 67.7 percentage points and increases ranged from 20 to 100 percentage points.

100% chance of completing college, only 166 also reported a 100% chance in 2001 and the average expectation of this group decreased by 21 percentage points.²⁴ In both cases, a large portion of students who reported at the two certainty points updated which suggests that at least at the extremes students are not Bayesian updaters. It is possible that students who reported either a 0% or 100% chance in the initial survey did so because they failed to understand the question or did not care enough to provide a reasonable answer. Therefore, further testing could reveal that students who report within the bounds of 0% and 100% are Bayesian updaters. Later tests utilize all of the 710 respondents first, then I limit the sample to only those who report between 1% and 99% in 1997.

Table 4.6: Certainty Priors

1997 Expectation	N	Number of Updaters	Percent that Update	Mean Change in Expectations
0%	25	11	44%	29.8
100%	265	99	37.4%	-21.0

Notes: The mean change in expectations is the average percentage point change in expectations for everyone who initially reported either a 0% or 100% chance of completing a four-year degree by 30. The number is larger if only those who update are included.

For those who did not report at the extremes I test whether new information is more valuable for those who were the most uncertain following the approach of Zafar (2013).²⁵ In the sample, the most value added should be for the students who reported near 50% and the least should be for those who report closer to the two bounds. I first tested this using a probit model where the dependent variable equaled 1 if the student updated in any way from 1997 to 2001 and 0 otherwise. I regressed this on an

²⁴The 99 respondents who decreased their expectations did so on average by 56.2 percentage points and the decreases ranged from -1 to -100 percentage points.

²⁵He does not test the extremes nor limits the sample in his testing, but concludes from his approach that students are Bayesian.

indicator for whether or not the individual reported between a 25% and 75% chance of completing college in 1997. If information is more valuable to these students they should be more likely than those who report in the tails to update their expectations. I then included this indicator in a model where the dependent variable is the absolute change in expectations. As a robustness check, I estimated a model with a similar dependent variable but included the square of 50 minus the individual's expectation in 1997 rather than the indicator. Both of these models are included below.²⁶

$$|\Delta Expectations_{2001-1997,i}| = \alpha + \beta(Expectation_{1997,i} \in [25, 75]) + \epsilon_i \quad (4.1)$$

$$|\Delta Expectations_{2001-1997,i}| = \gamma + \theta(50 - Expectation_{1997,i})^2 + \varphi_i \quad (4.2)$$

In the table below, column 1 contains the estimates from the probit model, columns 2 - 4 contain the estimates from equation 4.1, and columns 5-7 contain the estimates from equation 4.2. For equation 4.1, if $\beta > 0$, it indicates that individuals who were more uncertain about completing college in the initial survey made greater absolute changes in their beliefs. When the full sample is used the positive statistically significant estimates confirm the hypothesis that more uncertain individuals are both more likely to update their expectations and have larger absolute changes in their expectations. For equation 4.2, a negative estimate of θ would indicate that those who were more uncertain about completing college made greater absolute changes in their beliefs. Again, when the full sample is used our hypothesis that more uncertain individuals update more is confirmed. For both equations, the estimates had the hypothesized sign but were only significant at the 10% level when the indicators for those who reported either a 0% or 100% chance were included or when the sample was limited to only those who reported between 1% and 99% in

²⁶Specifications that included an indicator for those who reported a 50% chance were also estimated and the results varied based on which sample was used. For these estimates email the author at gray.hunter@uky.edu.

1997.

Table 4.7: Uncertainty Estimates

	Probit	Y= $ \Delta \text{ Expectations}_i $					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\text{Expect}_i \in [25, 75]$	0.921*** (0.120)	8.491*** (2.106)	5.071 (2.657)	5.071 (2.656)			
$(50 - \text{Expectation}_{1997,i})^2$					-0.00460*** (0.00106)	-0.00351 (0.00185)	-0.00351 (0.00185)
$\text{Expectations}_{1997} = 0$			2.432 (8.214)			5.301 (8.708)	
$\text{Expectations}_{1997} = 100$			-6.360* (3.097)			-3.491 (4.238)	
Observations	710	710	710	420	710	710	420
Adjusted R^2		0.016	0.021	0.007	0.021	0.022	0.008

Notes: Columns (3) and (7) contain estimates using only those who reported between a 1% and 99% chance initially.

- * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

- Robust standard errors in parenthesis

A Bayesian model also predicts differences in updating based on both how surprising the evidence is and the mass of the new evidence. The more surprising the evidence is to the individual the larger role it should play in the updating of the posterior expectation. I indirectly examine this by comparing how aligned the student's prior was with the estimated probability that they received the message they did. Students who reported high priors and were estimated to have a high probability of receiving positive messages in regards to college completion should not have been surprised. However, those who reported high priors, but were estimated to receive negative messages, or a lower likelihood of receiving a positive message, should have been more surprised, been more likely to update, and update in a larger fashion. The approach for estimating the likelihood of receiving a certain message is discussed in the section below. The greater the "mass" of the new evidence, or the weight that the

student places on that evidence in terms of it predicting college success, the larger the role it should play in the updating of their expectations. I am unable to estimate this as it would require a more in-depth interview with the respondent regarding what they believed to be the most significant predictors of college success and what exactly went into the formation of their expectations. Therefore, the remainder of the paper will focus on identifying whether or not students are Bayesian with an emphasis on if updating depends on the level of surprise of the messages received.

4.5 Estimating the Updating Process of Student's Expectations

Prior to estimating the informational component of the Bayesian model I first estimate a linear model to identify which sources of information discussed above potentially play a role in how or if the individual updates their expectation of completing college. I first estimate the equation below separately for each individual piece of new information and then include them all in one model. For a change in grades from 8th grade to high school I have created eight variables that capture any of the possible changes. For example, the Change As variable equals the difference between the indicator that the student received As in 8th grade and received As in High School or $Change\ As = (As\ in\ 8th\ grade - As\ in\ High\ School)$. The value of these variables range between -1 and 1 based on the grades they reported for 8th grade and high school so a separate matrix is need to interpret the coefficient and is included and discussed.

Linear Regression:

$$Expectation_{2001} = \gamma Information_{2001-1997} + \theta Expectation_{1997} + \epsilon_i$$

The estimates are obtained from using the full and restricted samples. The results

are similar so I focus on those from the restricted sample.²⁷ Across each specification the individual's prior plays a large role in the formation of the posterior which is in-line with what a Bayesian model predicts. Based on the assumptions and the specification the prior should be capturing all of the previous information that individuals were using to form their initial expectations and changes should be the result of only new information acquired. Examining the estimates from the new sources of information, a student's ASVAB score is positively associated with the posterior expectations.²⁸ This indicates that, holding their prior and other sources of information constant, those who score higher on the ASVAB are more likely to increase their expectation of completing college.

While students might be using their scores from the exam to update their expectations, it is more likely that both the lower and higher ability students are learning more about their ability and how it relates to completing college. It is likely they are learning through several sources of information and then incorporating it into their new expectations.²⁹ This is confirmed when examining the model where each of the new sources of information are included, estimates from which are in column 6 in the table below. When all of the information is controlled for, the magnitude of the ASVAB percentile decreases substantially but remains statistically significant which suggests a combination of the two hypothesis above.

²⁷Estimates from the full sample can be found in the appendix.

²⁸As a robustness check, I estimated the same model using only those who had an imputed ASVAB percentile. Since these students never received their ASVAB scores it should be insignificant if the scores were truly the information that students were using. It did still show up as significant, but the magnitude was much smaller. This suggests that higher ability students might report lower expectations initially, but once they learn their true ability to complete college through other sources they update their expectations accordingly.

²⁹Students do not receive their overall percentile score, but rather they receive their scores on the several different subject tests that the ASVAB is composed of.

Table 4.8: 2001 Expectation Formation - Interior Reporters

	Y = Expectation of completing a college degree in 2001					
	(1)	(2)	(3)	(4)	(5)	(6)
College Expectations 1997	0.540*** (0.0588)	0.794*** (0.0377)	0.668*** (0.0358)	0.849*** (0.0282)	0.523*** (0.0611)	0.478*** (0.0607)
ASVAB Percentile	0.507*** (0.0723)				0.487*** (0.0756)	0.335*** (0.0757)
Took the ACT or SAT Ever		11.95** (4.160)			3.129 (4.625)	-1.356 (4.344)
Enrolled in 2yr College			33.47*** (5.702)			31.05*** (6.142)
Enrolled in 4yr College			41.53*** (3.332)			33.17*** (4.271)
Change As				-1.917 (13.13)	-9.158 (12.73)	-9.602 (10.37)
Change As & Bs				-14.38 (12.78)	-19.38 (12.36)	-16.00 (10.31)
Change Bs				-17.20 (13.12)	-20.78 (12.55)	-19.05 (10.35)
Change Bs & Cs				-9.285 (12.92)	-18.06 (12.16)	-16.49 (10.04)
Change Cs				-6.669 (13.22)	-13.32 (12.70)	-11.01 (10.91)
Change Cs & Ds				-13.30 (13.65)	-21.16 (12.84)	-18.95 (11.12)
Change Ds				-5.271 (15.99)	-9.445 (15.66)	-6.983 (13.53)
Change Ds & other				0.246 (18.76)	-5.167 (17.99)	-0.904 (16.19)
Observations	341	420	420	420	341	341
Adjusted R^2	0.748	0.701	0.750	0.697	0.749	0.783

Notes: Only those who reported between 1% and 99% in the initial round were used to obtain these estimates. When the ASVAB percentile is used only those without imputed ASVAB scores were utilized.

- Robust standard errors in parenthesis

-* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The largest impact comes from the enrollment information indicators. In each case, if a student is enrolled in either a two or four-year college in 2001, their posterior expectations are much higher than their prior compared to those who are not enrolled at a post-secondary institution. This follows intuition as those who get into and attend a post-secondary secondary institution are more likely to go on to complete a degree

and believe they can do so. When run separately, whether or not a student took either the ACT or SAT exam was estimated to have a large impact on the student's posterior expectations, but when the other new sources of information are included in the model the impact is no longer statistically significant which supports the selection issue hypothesis. Last, it appears that changes in reported grades between 8th grade and high school do not significantly impact the respondents' posterior expectation per both an individual and joint test of significance. This is not surprising as the majority of students had already experienced high school when they were asked to report their initial expectations, but they were not asked to report their high school grades until the following year.³⁰ The lack of significance suggest that either students do not use grade changes to update their belief about their probability of completing college or that the high school grades are already being captured by the prior expectation.

³⁰A table that includes how students reported grades changed from 8th grade to high school as well as the predicted change in expectations for each change is included in the appendix.

Bayesian Regression Model with the ASVAB Message:

Below is the Bayesian model that is estimated:

$$Prob(S_{c,2001}|ASVAB_i) = \theta \left[\frac{Prob(ASVAB_i|S_c)}{Prob(ASVAB_i)} x Prob(S_{c,1997}) \right] + \epsilon_i$$

Posterior Expectation:

$Prob(S_{c,2001}|ASVAB_i)$ = the conditional (posterior) probability of completing college by 30 reported in 2001, given they received a certain ASVAB score.

The Impact and Processing of New Information:

$Prob(ASVAB_i|S_c)$ = the conditional probability of receiving a certain ASVAB score, given the true state is that they go on to complete college by 30.

$Prob(ASVAB_i)$ = the unconditional probability of receiving a certain ASVAB score.

Prior Expectation:

$Prob(S_{c,1997})$ = the unconditional (prior) probability of completing college by 30 reported in 1997.

For the 710 who reported in both periods, their expectation in 1997 is the prior and their expectation in 2001 is the posterior. I use the entire sample of 8,952 NLSY97 respondents to estimate both the unconditional and conditional probability of receiving each ASVAB score. I calculated the unconditional probability of receiving each of the 101 possible ASVAB scores by dividing the sum of the individuals that

received each of the 101 scores by the number of individuals who took the ASVAB.³¹ The conditional probability of each of the ASVAB scores given the true state is that the individual goes on to complete a four-year college degree was calculated by dividing the number of college graduates who received each score by the total number of college graduates. As an example, 80 students scored in the 25th percentile on the ASVAB and 12 went on to complete college. For someone who scored in the 25th percentile, their unconditional probability of receiving that score was $(\frac{80}{8,952})$, or .527%. Their conditional probability of receiving that score was $(\frac{12}{2,279})$, or .894%, where 2,279 is the number of individuals in the survey who had completed a four-year degree by 2013. The information and processing portion of the model is then just the conditional probability divided by unconditional probability.³² This estimate was assigned to each of the 710 students who answered in both periods based on their ASVAB score. Summary statistics for three primary component of the Bayesian model are included in table below.

Table 4.9: Summary Statistics of Bayesian Components

	Variable	Mean	Std. Dev.	Min	Max	N
Full Sample	$Prob(S_{c,2001} ASVAB_i)$	0.649	0.399	0	1	710
	$\frac{Prob(ASVAB_i S_c)}{Prob(ASVAB_i)}$	1.038	0.802	0	3.633	710
	$Prob(S_{c,1997})$	0.743	0.308	0	1	710
Interior Sample	$Prob(S_{c,2001} ASVAB_i)$	0.582	0.399	0	1	420
	$\frac{Prob(ASVAB_i S_c)}{Prob(ASVAB_i)}$	0.992	0.786	0	3.633	420
	$Prob(S_{c,1997})$	0.625	0.268	0	1	420

Notes: The interior sample are those who reported between 1% and 99% in 1997. See the paper for how each were calculated using the full NLSY97 sample.

³¹I rounded each of the respondent's ASVAB percentile scores to the nearest whole number so there are 101 possible scores between 0 and 100.

³²A table that displays the calculations of all three measures for each of the 101 ASVAB score possibilities is included in the appendix.

In the model, if students are Bayesian updaters then θ should be approximately equal to 1. This would indicate that the posterior expectation of completing college is a function of both old and new information. The prior expectation should contain all of the previous information students were using. The ASVAB component should include the new information students received from the ASVAB and how they believed it related to their ability to complete a college degree. The first column contains the estimate from the model using the full sample. A hypothesis tests indicates that the coefficient estimate is greater than zero, but not equal to one. As we showed prior, this could be the result of those who reported at the bounds of 0% and 100% who do not appear to update in a Bayesian fashion. To control for this, I estimate one model using the full sample with indicators for those who reported either a 0% or 100% chance and one that included only who reported between 1% and 99% initially. In both cases the estimate remains statistically greater than 0 but not equal to one.

Table 4.10: Bayesian Updating Estimates

	<u>Full Sample</u>	<u>Interior Sample</u>	<u>Full Sample</u>
	(1)	(2)	(3)
$\hat{\theta}$	0.503 (0.0156)	0.517 (0.0251)	0.398 (0.0178)
Expectation = 0 in 1997			0.298 (0.0790)
Expectation = 100 in 1997			0.331 (0.0310)
Observations	710	420	710
Adjusted R^2	0.592	0.549	0.641

Notes: Significance stars are for whether or not θ is = 1

-Robust standard errors in parentheses

- * $p < 0.05$, ** $p < 0.01$, ***

I pose that the estimates are a result of two possibilities. First, it might be that

students do not update in a Bayesian fashion. However, the positive statistically significant estimates indicate that students are updating based upon the message they received from the ASVAB but it potentially is not the only source they are using. Enrollment at either a two or four-year college also appeared to play a role in the posterior expectation but the endogenous nature of enrollment limits the analysis and scope of a measure that utilizes it. Future research will combat this issue as well as identify and test other potential sources.

4.6 Conclusions and Future Research

Obtaining reliable estimates of the returns to education requires that researchers successfully deal with the issues of selection. A popular approach is to model school choice using observed outcome data, assume that students have rational expectations, and that they condition their beliefs on similar variables process information in a similar way (Manski, 1993). However, students make schooling choices under uncertainty - uncertainty about personal tastes, individual abilities, and realizations of choice-related outcomes (Zafar, 2011). While the decision process has been examined theoretically (Manski, 1989; Altonji, 1993; Hilmer, 1998; Malumud, 2005) and empirically (Bamberger, 1987; Arcidiacono, 2004; Strange, 2012; Hunter, 2017a), most rely implicitly or explicitly on the unrealistic assumptions noted above regarding the expectations of students. If these assumptions are violated then a number of issues arise. First, if we are unaware of how youth perceive the returns to schooling, then it is impossible to infer the decision process from their observed schooling choices. Second, when not knowing a youth's decision process, inferring the objective returns to schooling from data on realized outcomes is implausible. These issues make an examination of how students form and update their expectations about their future education essential.

This paper has two goals, the first is to identify whether or not students incor-

porate the factors associated with college success into their expectation that they will complete college. The second is to test whether or not students update their educational expectations in a Bayesian manner. A unique question regarding the expectation that participants will go on to complete a four-year college degree, along with detailed responses from both the participants and their parents, from the 1997 National Longitudinal Survey of Youth (NLSY97) are used to investigate both topics. It is a nationally representative data set following individuals over time. Unlike previous studies, I observe students first while they are in high school and then follow them on an annual basis. This allows me to observe how their expectations change as they age and advance out of high school. I also am able to observe whether they fulfill their expectations. Their outcomes also are used to estimate whether or not they are Bayesian.

I begin by estimating which factors students appear to be utilizing when forming their educational expectations. I do this first for a subset of individuals in the NLSY97 that were between 15 and 17 years old when asked, and then later for another subset who were between 17 and 22 years old. I then compare the information that students appear to be using to form their educational expectations to factors that have been previously shown to impact college completion. I find that while high-school aged students are taking into account several of the relevant factors when forming their expectations, they also appear to be neglecting the impact that income and ability have on their likelihood of completing college. Overall, students are extremely overconfident, or overly aspirational, in their ability to complete a four-year college degree by the time they are 30 possibly because they are not incorporating these factors.

To test whether or not students are Bayesian, I use both samples mentioned above as well as 710 students who reported their expectations in the first round (1997) and then again 5 years later in the fifth round (2001). I develop a Bayesian model and discuss three predictions for how students should respond to the acquisition of

new information. I identify four sources of information that became available to students between 1997 and 2001 and test whether or not students responded to the new information in a way that the model would predict. I show that while students do appear to be utilizing the new information when updating their expectations, they do not appear to be true Bayesian updaters. I hypothesize two possibilities as to why this is not the case.

There are several avenues of future research that must be pursued. First, we must identify other potential sources of information that students might have and be using when forming their expectations. We must also examine whether or not students are updating their expectations rationally and how they relate to the observed outcomes. The next step is to compare the difference in information that respondents might have and identify if some information is more useful in helping students more accurately form their future expectations of academic success. If we can identify what information makes students "better" predictors of future events it might be possible to supply them with this earlier in life and allow them to make a more informed decision when deciding whether or not they should pursue a four-year college degree.

Chapter 5 Student Predictions and Alignment of Future College Success

5.1 Introduction

While increases in the published cost of college have been partially offset by additional grants or other forms of aid, the out-of-pocket cost of attending college has increased substantially over the past decade. In 2016-17, the estimated average net tuition and fees paid by full-time in-state students at public four-year institutions was \$3,770 higher than the net price a decade earlier and \$1,550 higher than the low of \$2,200 in 2009-10.¹ This cost increase and the associated student debt, coupled with worrisome reports of rising unemployment or under-employment by recent college graduates has people reassessing whether or not college is worthwhile investment.² However, what few have examined are the factors associated with the success or failure of students who decide to attend college. This is surprising as the 6-year graduation rate for first-time, full-time undergraduate students who began seeking a bachelor's degree at a four-year degree-granting institution in the Fall of 2009 was 59 percent.³ The rising out-of-pocket cost of college makes it more detrimental for students who attend college but are unable to complete the degree. They drop out and all they have to show for it is hundreds or even thousands of dollars in student loans. This paper is one of the few that examines students and the accuracy of their predictions that they will complete a four-year college degree by the time they are 30. The goal is to identify if the accuracy of predictions vary based on observable characteristics, and if so how this information can be used to help students form more

¹Prices vary considerably based on public vs. private, for-profit vs. non-profit, and whether or not the student is in-state or out-of-state but increases are consistent across each. CollegeBoard (2016)

²Kroeger et al. (2016)

³Greenstone and Looney (2012); Department of Education (2013); Shapiro et al. (2015)

accurate expectations of completing college. This could serve as one way to increase the graduation rates at post-secondary universities and reduce the percentage of those who take on excessive student loans and drop out.

The National Longitudinal Survey of 1997 (NLSY97) allows for a comparison of how a student's subjective probability of completing a four-year college degree reported while they are in high school and college compares to whether or not they are able to complete the degree by their late 20s to early 30s. The survey asks two subsets of students at two different times to report the probability they will obtain a four-year college degree by the time they turn 30. It then follows respondents on an annual basis up until 2013 when the majority of them are 30 or older which allows for an analysis of whether or not their previous expectations were fulfilled.⁴ It also allows for an analysis of how aligned students were initially and to examine and identify the similarities and differences between successful and unsuccessful predictors of college completion. Only one such study by Jacob and Wilder (2011) currently exists and they are limited by Likert scale responses by students and data that does not follow students on an annual basis.

I am able to use the large representative sample of the NLSY97 to compute the probability that a student will complete a four-year college degree by 2013 based on characteristics identified in the previous literature as being predictive of college success. This calculated probability is then used with the reported expectation for each of the students in a subset of the data to create two measures of alignment. I use these measures to examine the following questions. First, are there any observable differences between those who are aligned and misaligned? Next, do respondents become more aligned as they age and progress out of high school? Last, are those who are more aligned at an early age more likely to reach their outcomes?

⁴The NLSY97 followed individuals on an annual basis up until 2011 and then began collecting information every other year. The 2013 data is the most recently released data from the BLS.

I find that although the majority of students are overconfident, or more aspirational, in their belief that they will complete a four-year college degree, there are considerable differences in alignment based on several observable characteristics. In most cases the alignment of a student's expectations differ based on their: parent's education, ASVAB percentile, school enrollment, and race; alignment also differs based on other factors depending upon the age of the respondent and when they were asked. Using the two sub-samples of different aged respondents I show that as students age and acquire more information their expectation of completing college becomes more aligned with their estimated probability of completion, but the majority remain overconfident. I confirm this by examining 700 students who are asked their expectations first in 1997 while in high school then again five years later when they either are in college or the workforce. Last I show that those who are more aligned in either direction with what a model of college completion predicts the more likely they are to eventually reach that outcome.

5.2 Unique Aspects of the NLSY97 Data Utilized

Students' high school transcript data was collected by the BLS directly from the schools in two waves for those who provided written permission. These data can be broken down into individual terms for which the student was in school and contains their courses taken, grades, and school related activities in which they were involved. These data are available for 6,232 or 69% of the individuals who participated in the first round of the survey. The low reporting level signals possible selection by either parents or students who chose to release their grades; the primary alignment measure calculated using these data confirm this notion. However, the response rates were much higher for grades in both 8th grade and high school, as well as current schooling enrollment. While Black et al. (2011) find measurement error issues for individuals reporting their highest level of education attainment, when comparing

the students' self-reported grades to their transcripts they were relatively precise. Multiple models of alignment were estimated that used these measures separately and slight differences were present which are discussed in the next section.

This research utilizes the full sample to estimate the probability that an individual completes a four-year college degree. The two subsets that are asked to report their educational expectations are then used for the alignment analysis. Seven hundred individuals were asked their expectations in both 1997 and 2001 and are briefly examined. The 1997 and 2001 subsets have a similar gender and race composition compared to the full sample. Each are nearly 51% male and 49% female, and are approximately 52% white. The 1997 sub-sample is 27% black while the 2001 subset is 25.4% black. The two are also 20.8% and 23% Hispanic respectively. When they are separated into the cross-sectional and over-sample categories, there is more variation from the full sample.⁵ A means test for both race and gender show that the two expectations samples are not statically different from the full sample. Summary statistics for the full sample, and the 1997 and 2001 sub-samples are included in table 5.1 below.

⁵See the data section for details on the cross-sectional and over-sample categories.

Table 5.1: Full and Sub-sample Summary Statistics

Variable	Full Sample	1997 Sample	2001 Sample
College Expectations 1997	n/a	72.8%	n/a
College Expectations 2001	n/a	n/a	67.2%
Complete 4yr Degree	0.251	0.252	0.271
Representative Sample	0.751	0.747	0.753
Age in 1997	14.307	15.788	14.231
Age in 2001	19.055	20.561	18.948
Age in 2013	30.875	32.358	30.83
Male	0.512	0.504	0.511
White	0.519	0.52	0.511
–White Male	0.269	0.263	0.265
–White Female	0.251	0.256	0.246
Black	0.269	0.273	0.261
–Black Male	0.135	0.132	0.13
–Black Female	0.135	0.141	0.131
Hispanic	0.212	0.207	0.228
–Hispanic Male	0.109	0.109	0.116
–Hispanic Female	0.103	0.099	0.113
Household Size	4.549	4.492	4.583
No Parents Present	0.062	0.068	0.061
One Parent Household	0.315	0.324	0.301
Two Parent Household	0.623	0.609	0.638
Both Parents Less HS	0.099	0.095	0.104
Both Parents HS Degree	0.25	0.254	0.221
One Parent Some College	0.317	0.315	0.336
Both Parents Some College	0.061	0.056	0.07
One Parent College Degree	0.244	.241	0.243
Both Parents College Degree	0.097	0.093	0.101
ASVAB Percentile	44.762	44.858	45.435
–Imputed ASVAB	0.21	0.216	0.192
Gross HH Income	\$ 45312.23	\$ 46192.59	\$ 45276.61
–Imputed HH Income	0.267	0.275	0.255
High School GPA	2.818	2.797	2.832
Rural	0.268	0.269	0.242
Urban	0.732	0.731	0.758
No MSA	0.176	0.177	0.189
MSA Not Central City	0.502	0.499	0.495
MSA Central City	0.322	0.324	0.314
Census North East	0.176	0.173	0.174
Census North Central	0.228	0.238	0.214
Census South	0.374	0.38	0.381
Census West	0.222	0.21	0.232
Observations	8,984	3,511	1,946

Notes: The NLSY97 is composed of a representative portion and an over-sample portion. The percentage of the respondents in the representative sample is reported in the table above. The age in 2013 is reported because it is when the majority of the sample reported their highest educational achievement. For the full sample and 1997 sub-sample, the location variables are what is reported in 1997; the 2001 reported location is used for the 2001 sub-sample. See the paper for notes on the imputation of both the ASVAB percentile and Household Income.

5.3 Estimation and Results

Predicting College Completion

The first step in creating the measures of alignment was to estimate the probability that an individual would complete college based on several characteristics known to be associated with college success. A probit model was run on the full sample and the coefficients were saved. Then for each individual in both sub-samples the probability that they would complete college based on their observed characteristics was calculated. For both sub-samples, controls for observables such as age, race, gender, ASVAB percentile, family income, household structure, and parent's education were included. Then, for the 1997 sub-sample two separate predictions were made. The first was created using the student's self-reported grades in 8th grade and the other used their high school GPA from their high school transcript.⁶ Next, since the 2001 sub-sample was older and should have had access to more information when asked their expectation, enrollment in high school, a two-year college, or a four-year college could be used. One model using the high school transcript GPA was estimated and another using these enrollment controls instead was estimated. A summary of the predictions from each model can be found in tables 5.2 and 5.3 and are discussed in detail below.

Testing the validity of the probit model predictions, both the full and sub-sample predictions are included in the tables as well as the mean completion rate for each sample. It is also further separated by those who go on to complete college by 2013 and those who do not to further analyze the measure's validity. Comparing the predictions from both of the models that used the high school transcript GPA the selection of those who consented to have their school release their GPA to the Bureau

⁶The students' self-reported high school grades were not used as they were not asked of the respondent until the second round.

of Labor Statistics is evident. For the full sample that used the 1997 controls the probit model predicted an average probability of completion of 29.6% and 29.58% of those who allowed access to their high school GPA went on to complete the degree by 2013. Note the limited sample size and higher completion rates when compared to the model that used the self-reported grades and the overall sample. Only around 25% of the NLSY97 sample had completed a four-year college degree by 2013 which suggests that this model suffers from selection.⁷ These statistics are similar for the model that uses the transcript GPA for the 2001 sub-sample. While the summary statistics for the models that used the transcript data are discussed, estimation will focus primarily on the models that used the student's self-reported grades in 1997 and the enrollment controls in 2001 as they appear to provide more accurate predictions of future college success.

Table 5.2: Probit Model Predictions of College Completion - 1997 Sub-sample

Controls Used to Predict Completion	Sample	N	Mean	Std. Deviation	Min	Max	Percent that Complete Degree by 2013
1997 Grades	Full Sample:	8865	0.254	0.255	0.000	0.976	25.14%
	Do Not Complete College	6606	0.167	0.185	0.000	0.976	0%
	Complete College	2259	0.510	0.260	0.006	0.976	100%
	1997 Sub-sample:	3475	0.257	0.255	0.000	0.976	25.21%
	Do Not Complete College	2590	0.170	0.185	0.000	0.949	0%
	Complete College	885	0.512	0.263	0.011	0.976	100%
1997 HS GPA	Full Sample:	6004	0.296	0.273	0.000	0.986	29.58%
	Do Not Complete College	4228	0.189	0.197	0.000	0.981	0%
	Complete College	1776	0.550	0.259	0.003	0.986	100%
	1997 Sub-sample:	2561	0.295	0.272	0.000	0.986	29.29%
	Do Not Complete College	1811	0.191	0.196	0.000	0.966	0%
	Complete College	750	0.548	0.264	0.006	0.986	100%

Notes: The mean reported is the average predicted probability of completing college for the sample utilized from a Probit model that included similar controls to essay 1 but excluded the student's expectations. The two portions represent different controls used in each model as there were concerns about selection based on who chose to report their High School GPA.

⁷25% is also the estimated percent of the population during that time that held a four-year college degree per the Bureau of Labor Statistics, U.S. Department of Labor (2013).

Table 5.3: Probit Model Predictions of College Completion - 2001 Sub-sample

Controls Used to Predict Completion	Sample	N	Mean	Std. Deviation	Min	Max	Percent that Complete Degree by 2013
2001 HS GPA	Full Sample:	5518	0.307	0.318	0.000	0.994	30.86%
	Do Not Complete College	3815	0.160	0.205	0.000	0.988	0%
	Complete College	1703	0.636	0.275	0.002	0.994	100%
	2001 Sub-sample:	1348	0.316	0.320	0.000	0.994	30.95%
	Do Not Complete College	930	0.165	0.214	0.000	0.988	0%
	Complete College	418	0.651	0.257	0.047	0.994	100%
2001 Enrollment	Full Sample:	7808	0.269	0.296	0.001	0.987	27.07%
	Do Not Complete College	5694	0.148	0.192	0.001	0.978	0%
	Complete College	2114	0.593	0.283	0.010	0.987	100%
	2001 Sub-sample:	1933	0.275	0.298	0.001	0.983	27.13%
	Do Not Complete College	1408	0.152	0.197	0.001	0.978	0%
	Complete College	525	0.606	0.270	0.013	0.983	100%

Notes: The mean reported is the average predicted probability of completing college for the sample utilized from a Probit model that included similar controls to essay 1 but excluded the student's expectations. The two portions represent different controls used in each model as there were concerns about selection based on who chose to report their High School GPA.

Measures of Alignment

I calculate two measures of alignment using the predicted probabilities of college completion above. The first is the difference between the individuals reported expectation of completing college by 30 and the predicted probability of completing the degree by 2013 discussed above. The second is the ratio of the individual's expectation and their predicted probability of completing college. Equation 5.1 below presents the first and primary measure of alignment and Equation 5.2 presents the secondary alignment measure. In both cases, i represents the individual and t indicates which sub-sample is used.

$$Primary\ Alignment_{i,t} = \left(Expectation_{i,t} - P(\widehat{Complete\ College}_{i,2013}) \right) \quad (5.1)$$

$$Secondary\ Alignment_{i,t} = \left(\frac{Expectation_{i,t}}{P(\widehat{Complete\ College}_{i,2013})} \right) \quad (5.2)$$

The construction of the primary measure bounds the values it can take between

-1 and 1, with zero indicating that the respondent's reported expectation is perfectly in-line with their predicted probability of completing college. Positive values indicate that the individual is overconfident, or more aspirational. Negative values indicate they are under-confident, or have lower aspirations. I use the measure to classify respondents in one of four ways. They can be under-confident, aligned, overconfident, or extremely overconfident. Under-confident respondents, which is the least populated group, are those whose alignment measure was between -1 and -.26. Aligned individuals are those whose measure was between -.25 and .25. Overconfident respondents measured between .26 and .75, and those who were extremely overconfident are those whose measure was than .75. A distribution of the primary alignment measure for each sub-sample conditional on the controls used in the prediction probit model are presented below.

Figure 5.1: Primary Alignment Measure for each Sub-sample

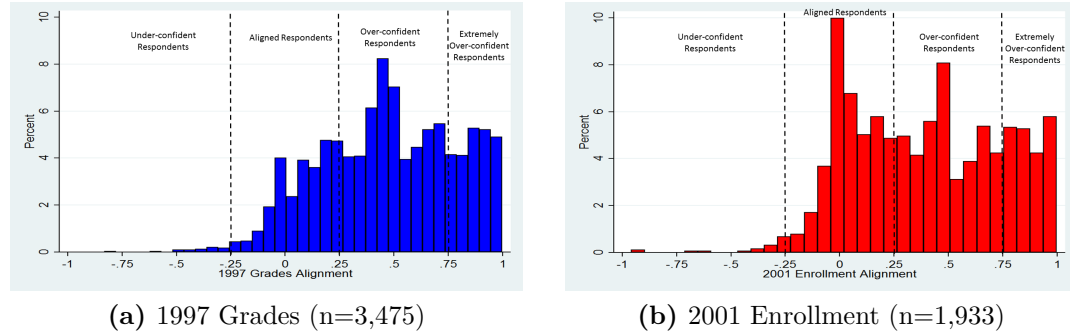
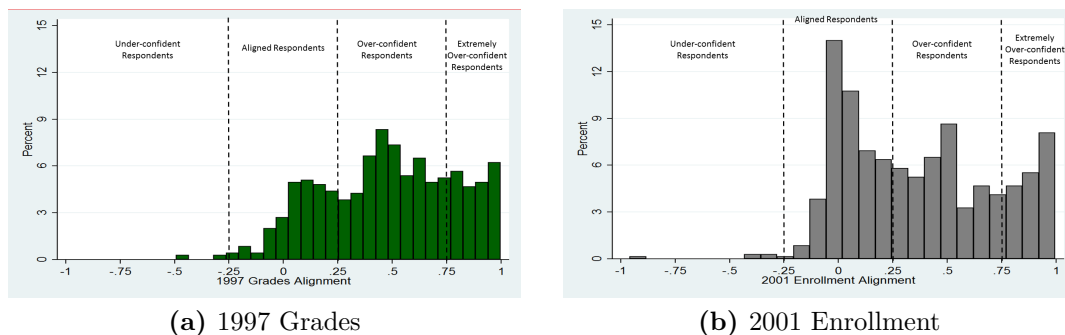


Figure a is for the 1997 sub-sample and includes the calculated measure that used the student's self-reported grades. The figure indicates that students between 15 and 17 years old are primarily overconfident in their ability to complete a four-year college degree. A number of students also appear to be aligned, but nearly none of them are under-confident. Figure b presents a similar measure for the 2001 sub-sample that includes the calculated measure that used the high school and college enrollment controls. Recall that these respondents were asked their expectations in the fifth

round of the survey and the age range was not restricted so this group included those between 17 and 22 so they should have more information about what is needed to complete a college degree. Therefore, if they are incorporating this additional information into their expectations then they should be more aligned as a group. The distribution suggest that this is the case as a larger percentage of students in this group are more aligned, and less are either overconfident or extremely overconfident. I confirm this by examining the distributions for the 700 respondents who answered the expectations question in both 1997 and 2001. While a portion of respondents remain overconfident, a much larger percentage are aligned in 2001 than in 1997. The distributions can be found below.⁸

Figure 5.2: Primary Alignment Measure using the Overlapping Sub-sample



Summary statistics for the secondary alignment measure are included in table 5.4. This measure is bounded between 0 and 100.⁹ A perfectly aligned individual is one whose expectation equaled the predicted probability that they would complete college. If this were the case, the secondary alignment measure would equal one. If their expectation was greater than the predicted probability the measure would be greater

⁸Summary statistics of the variation in the primary alignment measure based on race, gender, and race and gender can be found in the appendix. Summary statistics for all variables, separated by each category of confidence noted above are also present in the appendix.

⁹Those who received an alignment measure of greater than 100 were excluded as these larger outliers were skewing the statistics. There were 69 such individuals in the 1997 sub-sample and 35 in the 2001 sub-sample.

than one, and if the expectation was lower it would be less than 1. Similar to the primary measure, the secondary measure indicates that the majority of the individuals were overly optimistic about their ability to complete college. I again categorize individuals as under-confident, aligned, overconfident, or extremely overconfident based on the alignment measure. The individual was classified as under-confident if their alignment measure was less than .75, aligned if their measure was between .75 and 1.25, overconfident if their measure was between 1.25 and 2, and extremely overconfident if their measure was greater than 2. Statistics for both the 1997 and 2001 sub-sample conditional on each of these classifications are presented.

Table 5.4: Summary Statistics for the Secondary Alignment Measure

Sample	N	Mean	Standard Deviation	Min	Max	Percent that Complete Degree by 2013
<u>1997 Sub-Sample</u>						
<u>Full:</u>	3406	8.87	13.65	0	98.10	25.98%
-Under-confident	232	0.10	0.21	0	0.74	3.88%
-Aligned	288	1.09	0.11	0.77	1.25	68.75%
-overconfident	576	1.58	0.21	1.25	2.00	52.95%
-Extremely overconfident	2310	12.54	15.26	2.00	98.10	16.15%
<u>2001 Sub-Sample</u>						
<u>Full:</u>	1898	8.43	14.85	0	97.52	27.66%
-Under-confident	301	0.04	0.14	0	0.75	1.66%
-Aligned	236	1.10	0.09	0.75	1.25	72.46%
-overconfident	288	1.55	0.20	1.25	2.00	62.85%
-Extremely overconfident	1073	14.25	17.66	2.00	97.52	15.66%

Notes: The secondary measure of alignment is the ration of the individual's reported expectations and the estimated probability that they will complete a college degree using the sample specific controls outlined in the paper. Those who received an alignment measure of over 100 are not included in these statistics.

Consistent with the findings from the primary measure, the statistics from the secondary alignment measure also suggest that as students age they updated their

expectations downward on average. In 1997, 6.8% of the sub-sample was under-confident, 8.5% were aligned, 16.9% were overconfident, and 67.8% were extremely overconfident. In 2001, 15.9% of the sub-sample was under-confident, 12.4% were aligned, 15.2% were overconfident, and 56.5% were extremely overconfident. A larger portion of the 2001 respondents were either under-confident or aligned and less were overconfident or extremely overconfident. This suggests that students are learning about their true probability of completing college as they age and are incorporating this into their expectations.

The last column includes the percent of each category that completed a four-year college degree by 2013. Regardless of sub-sample, those who were aligned went on to complete college at a much higher rate than those in any of the other three categories. Those who were overconfident also completed college at a much higher rate than the sample average, but those who were extremely overconfident completed college at rate around 10 percentage points lower than the sample average. This suggests that the alignment of expectations is associated with future college success and confidence can be beneficial to a point. Further analysis is needed as the measure does not distinguish between those who have high expectations and a high probability of completing college and those who have low expectations and a low probability of completing college.¹⁰ From both measures we find that students on average are overly optimistic in their ability to complete a four-year college degree but their alignment does appear to be increasing as they age. There also appears to be a positive relationship between alignment and college completion. The remainder of the paper will utilize only the primary alignment measure as the findings are more easily interpretable.¹¹

¹⁰This should bias the current completion statistics downward as those who did not complete college that had low expectations and a low probability of completing are accurately predicting that they will not complete a college degree.

¹¹Later versions will include an analysis of the secondary measure.

Examining the Variation in Alignment

To examine how student alignment varies a simple linear model is estimated with the dependent variable being the square of the primary alignment measure. The squared variable is used because few students are under-confident and the estimated coefficients are more easily interpretable. A positive estimate indicates that the variable is negatively related to student alignment, while a negative estimate indicates a positive relationship with student alignment.¹² I hypothesize that alignment potentially differs based on various characteristics such as demographics, area where the respondent lived, household structure and composition, parental education, family income, and measures of previous school success so controls for each are included in the model. In addition to these, when the 2001 alignment measure is used I also include controls for their current enrollment in high school, a two-year college, or a four-year college. The estimates from each can be found in the tables below.

As discussed above, an individual can be aligned in several ways but I focus on the two most common. The first are those who report a low expectation and who have a low estimated probability of completion. They are aligned in that they predicted they would not complete college and the data suggested that, based on their observables, that this was the case. The second, and most prevalent in the data, are those who report a high expectation and have a high estimated probability of completion. The model below does not differentiate between these two forms or any of the other forms of alignment. Therefore, when discussing the estimates I condition my hypotheses concerning the significance, sign, and magnitude of each on the possible forms of alignment. Further discussions of the types of alignment and realized outcomes are included in a later section.

¹²A model where all of the under-confident individuals were excluded was also estimated and the results did not noticeably differ. For these estimates email the author.

Table 5.5 includes estimates using the full 1997 sub-sample in column 1. The estimates indicate the following variables are statistically related to student alignment: age; white and Hispanic; if there are 1 or 2 parents present in the household; if the parents both have less than a high school degree or if they have at least some college experience; ASVAB percentile; household income quantile; good grades in 8th grade and enrollment in high school; and if the student lives outside of the central city but inside and MSA. All of these factors, excluding being enrolled in high school at the time of the survey, are estimated to increase the likelihood that the student is aligned. This is likely a result of the lack of differentiation between the two most common types of alignment, and most can be justified for improving alignment in either direction.

I discuss how these factors are related to alignment in terms of how they potentially relate to information that the student might or might not have. The positive impact of age is expected as older students should be gathering more relevant information in terms of their ability to complete college and incorporating this into their expectations.¹³ The data also support the idea that parents play a valuable role in the information that their children have regarding college. First, students who are being raised by at least one parent are more aligned than those who are raised in a household without one. Next, those who had biological parents who both did not complete high school or who had at least one parent with some college experience were more aligned than those who had parents with high school degrees. This again is likely a result of the various forms of alignment. Whether or not one or both parents were in the household combined with their educational experiences are sources of information that the child utilizes to form their expectation of completing college in the future. When they lack these sources their expectations appear to be less precise.

¹³See essay 2 above that supports these findings

Table 5.5: Estimates of Expectation Alignment in 1997

	$Y = (Expectation_{i,1997} - P(CompleteCollege_{i,2013}))^2$	
	1997 Sub-sample	Overlapping 700
Age in 1997	-0.0173** (0.00632)	-0.0169 (0.0144)
Male	-0.00223 (0.00912)	0.00296 (0.0205)
White	-0.0591*** (0.0149)	-0.0703* (0.0351)
Hispanic	-0.0423** (0.0156)	-0.0414 (0.0344)
One Parent HH	-0.0704** (0.0224)	-0.127* (0.0530)
Two Parent HH	-0.0563** (0.0211)	-0.0954 (0.0490)
Both Parents Less than HS	-0.0628** (0.0212)	-0.0344 (0.0469)
Both Parents HS Degree	-0.0221 (0.0152)	-0.0257 (0.0343)
One Parent Some College	-0.00290 (0.0133)	-0.0196 (0.0282)
Both Parents Some College	-0.0509** (0.0171)	0.000296 (0.0333)
One Parent College Degree	-0.0396** (0.0129)	-0.0358 (0.0276)
Both Parents College Degree	-0.0956*** (0.0141)	-0.0901** (0.0316)
ASVAB Percentile	-0.00209*** (0.000213)	-0.00230*** (0.000483)
Gross HH Income Quantile	-0.0139** (0.00478)	-0.0120 (0.00996)
As 8th Grade	-0.156*** (0.0172)	-0.117** (0.0361)
As & Bs 8th Grade	-0.0630*** (0.0170)	-0.0165 (0.0357)
Bs 8th Grade	0.00210 (0.0186)	0.0475 (0.0393)
Bs & Cs 8th Grade	0.0328 (0.0180)	0.0626 (0.0391)
Cs & Ds 8th Grade	-0.0372 (0.0253)	-0.0310 (0.0599)
Ds 8th Grade	-0.0346 (0.0388)	0.0378 (0.0756)
Enrolled in HS in 1997	0.163*** (0.0243)	0.130* (0.0588)
Urban in 1997	0.0191 (0.0118)	-0.00491 (0.0263)
In MSA, Not Central City in 1997	-0.0313* (0.0135)	0.00984 (0.0303)
In Central City MSA in 1997	-0.0186 (0.0161)	0.0356 (0.0351)
North East Census in 1997	-0.0126 (0.0129)	-0.0169 (0.0284)
North Central Census in 1997	-0.00698 (0.0117)	-0.00172 (0.0271)
West Census in 1997	-0.0142 (0.0129)	-0.0624* (0.0281)
Observations	3475	708
Adjusted R^2	0.215	0.227

Notes: The model included a statistically significant constant with a coefficient estimate of approximately .75 but was not included. Each of the specifications above was also estimated without a constant and the majority of the estimates excluding age were similar in sign and significance.

- Robust standard errors in parentheses

- * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The ASVAB and academic success measures also suggest that ability and previous academic success play a role in forming accurate expectations. Those who scored higher on the ASVAB also reported more aligned expectations, as well as those who

reported earning mostly As and Bs or higher in the 8th grade.¹⁴ We expect that those who received lower grades would also be more aligned in terms of not completing college, but the data suggest otherwise. Those who reported Cs in 8th grade are similarly aligned to those who reported grades below that.¹⁵ The negative impact that enrollment in high school has on student alignment seems counter-intuitive, but those who were not enrolled in 1997 also reported low expectations of completing college and few ultimately completed a four-year degree. This indicates that those who dropped out did so knowing they would more than likely not complete college and reported their expectations as such.

The most surprising and concerning of the estimates are those regarding race. They suggest that both whites and Hispanics are considerably more aligned at a young age than blacks even after controlling for age, ability, family characteristics, and location. The data indicate that blacks are extremely overconfident in their ability to complete a college degree, especially when they are in high school.¹⁶ It is possible that blacks simply paid less attention during the interview and on average reported much higher expectations as a way to hurry the survey. They might also be surrounded by more people telling them that they need to go to college and thus have a higher belief of being able to complete college than others. Regardless, why blacks have more misaligned expectations of completing college than their white and Hispanic counterparts must be the focal point of future research.

A similar model of alignment was estimated using the 2001 sub-sample, the estimates of which are in column 1 of table 5.6. Recall that this sub-sample includes individuals between 17 and 22 years old when they reported their expectations in

¹⁴Although the majority of these students were in high school in 1997, they were not asked to report their high school grades until the second round so they are not included in this model.

¹⁵A large portion of the 1997 sub-sample was enrolled in high school at the time of the survey, so it is possible individuals are using their most recent grades as predictors rather than those from the past.

¹⁶These findings are in-line with Hoelter (1982) regarding the rationality of aspiration formation.

2001, and were on average 3 years older than those in the 1997 sub-sample. Therefore, comparing estimates from the two could provide valuable information regarding how students form their expectations as they age. From the 2001 model, parental education is again significantly related to student alignment but in a more limited fashion. Only those who had parents who both did not complete high school or both who completed a college degree were more likely to be aligned than those who had parents who held a high school degree. Whether or not one or both parents were present in the household is no longer associated with aligned expectations. These together suggest that as students age and transition out of high school they rely more on what they have learned about themselves and the true costs of college, and less on their parents when forming their expectations. This is confirmed by the significance of the control for enrollment at a four-year college. Students enrolled at four-year college were more likely to report aligned expectations than those who were not enrolled in any form of schooling and those who were enrolled at a two-year college.

The association between high school grades and alignment implies that students might be over-weighting the value of success in high school and how it relates to college success. Those who reported earning As and Bs in high-school were on average less aligned than those who reported mainly Cs. It is possible that students believe marginal success in high school will translate to success in college, but that might not be the case with recent issues of grade inflation and findings that a large portion of students who get to college are academically under-prepared.¹⁷ Those who reported receiving Cs and Ds or below are on average more aligned than those who reported Cs. I again suggest that this is because they take it as a signal that they lack the ability to complete a college degree and report expectations that are in-line with this. This also supports my hypothesis that students rely more heavily on their most recent

¹⁷See Bettinger and Long (2009)

grades when forming their expectations, and justifies the insignificance of the controls for lower grades while in 8th grade in the model that used the 1997 sub-sample. Those who reported living in an urban area or in the North East Census region in 2001 were also estimated to be less aligned but endogenous migration is a concern.

Estimates from the 2001 sub-sample again suggests that whites have more aligned expectations than black, but there is no longer a statistically significant difference between Hispanics and blacks. The lack of difference between blacks and Hispanics in the older sub-sample is encouraging, but the difference between blacks and whites persists and indicates potential differences in how whites and blacks both form and update their expectations of completing college. These differences appear to result in blacks forming more misaligned prospects about themselves completing college and could result in issues with the decision to attend and/or complete college. A more detailed analysis as to whether this is a result of under or over-confidence is needed, as well as how this might be impacting college choice, persistence, and completion of college for those of different races.¹⁸

¹⁸The NLSY97 data indicates it is more than likely a result of over-confidence especially for black males.

Table 5.6: Estimates of Expectation Alignment in 2001

	$Y = (Expectation_{i,2001} - P(CompleteCollege_{i,2013}))^2$	
	2001 Sub-sample	Overlapping 700
Age in 2001	-0.00214 (0.00612)	-0.00124 (0.0179)
Male	0.0171 (0.0129)	-0.00514 (0.0220)
White	-0.0619** (0.0214)	-0.0697 (0.0386)
Hispanic	-0.0324 (0.0225)	-0.0493 (0.0385)
One Parent HH	-0.0452 (0.0325)	0.0226 (0.0511)
Two Parent HH	-0.0331 (0.0311)	0.0472 (0.0459)
Both Parents Less than HS	-0.0698* (0.0278)	-0.0772 (0.0467)
Both Parents HS Degree	-0.0251 (0.0219)	-0.0537 (0.0359)
One Parent Some College	-0.00292 (0.0190)	-0.0524 (0.0319)
Both Parents Some College	0.00742 (0.0247)	0.0177 (0.0453)
One Parent College Degree	-0.0134 (0.0189)	-0.00709 (0.0319)
Both Parents College Degree	-0.0851*** (0.0225)	-0.0957* (0.0393)
ASVAB Percentile	-0.00116*** (0.000309)	-0.000110 (0.000556)
Gross HH Income Quantile	-0.00617 (0.00656)	-0.00907 (0.0111)
As HS	0.0284 (0.0249)	0.0132 (0.0424)
As & Bs HS	0.0797*** (0.0232)	0.107** (0.0411)
Bs HS	0.0331 (0.0237)	0.0498 (0.0402)
Bs & Cs HS	0.0227 (0.0227)	0.0442 (0.0377)
Cs & Ds HS	-0.0595* (0.0300)	-0.0203 (0.0524)
Ds HS	-0.123* (0.0586)	0.00161 (0.116)
Below Ds in HS	-0.209*** (0.0452)	-0.161** (0.0491)
Enrolled in HS in 2001	0.00544 (0.0217)	0.0345 (0.145)
Enrolled in a 2yr college in 2001	0.0458 (0.0235)	0.0507 (0.0336)
Enrolled in a 4yr college in 2001	-0.200*** (0.0157)	-0.199*** (0.0230)
Urban in 2001	0.0421* (0.0169)	0.0619* (0.0302)
In MSA, Not Central City in 2001	0.0193 (0.0181)	0.0252 (0.0313)
In Central City MSA in 2001	0.0300 (0.0216)	0.00457 (0.0366)
North East Census in 2001	0.0410* (0.0190)	0.00906 (0.0332)
North Central Census in 2001	0.0254 (0.0168)	0.00914 (0.0291)
West Census in 2001	-0.0169 (0.0180)	-0.0305 (0.0318)
Observations	1933	707
Adjusted R^2	0.171	0.161

Notes: The model included a statistically significant constant with a coefficient estimate of approximately .44 for the full model but was not included; the constant was not statistically significant for the overlapping 700. Each of the specifications above was also estimated without a constant and the majority of the estimates excluding age and enrollment in high school or a two-year college were similar in sign and significance.

- Robust standard errors in parentheses

- * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Next I compare the estimates for each model that utilized only the 700 individuals who reported their expectations in both 1997 and 2001. The estimates from each are in column 2 of their respective tables above. In both cases the sign and magnitude of the coefficients were similar to the models that used the full sub-samples, but the limited sample size and obvious multicollinearity resulted in larger standard errors and decreased the significance of the estimates. When these 700 students were younger, there were estimated differences in alignment based on race, whether parents were present in the household, if both of their parents had a college degree, ASVAB percentile, receiving As in 8th grade, and enrollment in high school. These are similar to when the full 1997 sub-sample was used. When the 700 reported their expectations five years later, the controls for parental education, college enrollment, and high school grades were all significant. Those with parents who both had at least a college degree, were enrolled in a four-year college, or received lower grades in high school had on average more aligned expectations. Reporting As and Bs in high school indicated having less aligned expectations. Whites, blacks, and Hispanics appear to be similarly aligned when the individuals reported the second time. This could be the result of only the oldest portion of the respondents in the 2001 sub-sample being utilized, which might indicate that blacks do become more aligned but at slower pace than whites and Hispanics. A more detailed analysis is needed to test this hypothesis.

Alignment and the Realization of One's Expectations

Above I've shown that there are several factors associated with individuals having more or less aligned expectations. However, neither of the alignment measures indicate whether or not the individual is aligned in terms of predicting future failure or success. Below I examine the relationship between the type of one's alignment and whether or not they fulfill their expectations. For each sub-sample, I separate and compare the expectations and completion rates of those who were aligned, overconfi-

dent, and extremely overconfident. Within each category I separate the respondents into one of several bins based on their expectation. Comparing the aligned, overconfident, and extremely-overconfident I again find a positive relationship between alignment and college completion. When I compare within each of the categories, I also find those who were aligned and reported low expectations rarely completed college while those who were aligned and reported higher expectations went on to complete college at high rate. Overconfident respondents that reported an expectation of 75% or higher completed at a higher rate than the sample, but those who were extremely overconfident graduated at a much lower rate.

Table 5.7: Expectation Alignment and Completion

1997 Sub-Sample				2001 Sub-sample		
	Expectations	N	% that Complete College Degree	Expectations	N	% that Complete College Degree
Aligned Respondents	0%	164	1.22%	0%	252	0.79%
	1-25%	271	1.48%	1-25%	177	0.56%
	26-49%	35	2.94%	26-49%	22	4.55%
	50%	81	23.46%	50%	18	16.67%
	51-74%	22	54.55%	51-74%	4	25.00%
	75-99%	167	77.84%	75-99%	77	76.62%
	100%	151	83.44%	100%	179	83.24%
		891	33.00%		729	29.63%
overconfident Respondents	26-49%	66	3.03%	26-49%	38	0.00%
	50%	431	3.71%	50%	173	0.58%
	51-74%	126	12.70%	51-74%	52	3.85%
	75-99%	622	28.14%	75-99%	220	33.18%
	100%	502	51.00%	100%	279	59.86%
		1,747	26.62%		762	31.89%
Extremely overconfident Respondents	75-99%	187	6.42%	75-99%	120	11.67%
	100%	565	15.04%	100%	270	15.19%
		752	12.90%		390	14.10%

Notes: Per the construction of the measure, the more confident individuals could not report below certain thresholds. This is the reason that 0 reported less than a 26% chance of completing college for the overconfident respondents. The same is true for under 75% for the extremely overconfident.

For the aligned respondents, on average those at both ends of the distribution

were more likely to realize their expectation. For both sub-samples, less than 2% of those who reported an expectation of 0% went on to complete college; the pattern is similar for those who reported an expectation of less than a 50% as no more than 5% of them completed college by 2013. As expectations increase so does the percentage of those who completed the degree. Over 76% of those in the aligned category that reported an expectation between 75% and 99% completed college; over 83% of those who reported a 100% chance completed college.

The positive relationship between expectations and completion persists for both the confident and under-confident individuals but the accuracy of the predictions diminishes significantly. Only 3% of those who were overconfident and reported an expectation less than 50% completed college which is similar to the aligned respondents. However, less than 4% of those who reported a 50% chance in either sub-sample completed college. Less than 33% of the overconfident respondents who reported an expectation between 75% and 99% went on to complete college compared to 76% for the aligned students who reported similar expectations. Last, 51% of those who reported an expectation of 100% in 1997 went on to complete college, while 59% of those who reported an expectations of 100% in 2001 completed college. The differences between the aligned and the extremely overconfident are even more drastic. Overall, less than 15% of those who were extremely overconfident and reported an expectation between 75% and 100% completed a college degree by 2013.

Overall, the alignment of one's expectations and the realization of the outcome appear to be directly linked. Those who reported expectations in-line with what a model of school completion predicts on average fulfill their expectations at a much higher rate than those who were overconfident. The positive relationship between alignment and college completion is driven by two factors. First over 75% of those who were aligned and reported an expectation between 75% and 100% completed college. Second, only 14% of those who were extremely overconfident and reported

an expectation within the same bounds went on to complete college. These seem to dominate the large portion of students who were aligned and fulfilled their low expectations by not completing college. The next step is to look into how we can improve the accuracy of students' expectations. The data show that the older, 2001 sub-sample contained a larger percentage of students who reported aligned expectations. In the 1997 sub-sample approximately 25% of students were aligned, while 39% of those in the 2001 sub-sample were aligned. This indicates that students are learning as they age and incorporating new sources of relevant information into their expectations. Future research is needed to identify the information that students are obtaining and using to form more aligned expectations as they age. If this can be identified, it could be used to help students form more accurate expectations of completing college prior choosing whether or not to attend college.

5.4 Conclusions

While increases in the published cost of college have been partially offset by additional grants or other forms of aid, the out-of-pocket cost of attending college has increased substantially over the past decade. In 2016-17, the estimated average net tuition and fees paid by full-time in-state students at public four-year institutions was \$3,770 higher than the net price a decade earlier and \$1,550 higher than the low of \$2,200 in 2009-10.¹⁹ This cost increase and the associated student debt, coupled with worrisome reports of rising unemployment or under-employment by recent college graduates has people reassessing whether or not college is worthwhile investment.²⁰ However, what few have examined are the factors associated with the success or failure of students who decide to attend college. This is surprising as the 6-year

¹⁹Prices vary considerably based on public vs. private, for-profit vs. non-profit, and whether or not the student is in-state or out-of-state but increases are consistent across each. CollegeBoard (2016)

²⁰Kroeger et al. (2016)

graduation rate for first-time, full-time undergraduate students who began seeking a bachelor's degree at a four-year degree-granting institution in the Fall of 2009 was 59 percent.²¹ The rising out-of-pocket cost of college makes it more detrimental for students who attend college but are unable to complete the degree. They drop out and all they have to show for it is hundreds or even thousands of dollars in student loans. This paper is one of the few that examines students and the accuracy of their predictions that they will complete a four-year college degree by the time they are 30. The goal is to identify if the accuracy of predictions vary based on observable characteristics, and if so how this information can be used to help students form more accurate expectations of completing college. This could serve as one way to increase the graduation rates at post-secondary universities and reduce the percentage of those who take on excessive student loans and drop out.

The National Longitudinal Survey of 1997 (NLSY97) allows for a comparison of how a student's subjective probability of completing a four-year college degree reported while they are in high school and college compares to whether or not they are able to complete the degree by their late 20s to early 30s. The survey asks two subsets of students at two different times to report the probability they will obtain a four-year college degree by the time they turn 30. It then follows respondents on an annual basis up until 2013 when the majority of them are 30 or older which allows for an analysis of whether or not their previous expectations were fulfilled.²² It also allows for an analysis of how aligned students were initially and to examine and identify the similarities and differences between successful and unsuccessful predictors of college completion. Only one such study by Jacob and Wilder (2011) currently exists and they are limited by Likert scale responses by students and data that does not follow

²¹Greenstone and Looney (2012); Department of Education (2013); Shapiro et al. (2015)

²²The NLSY97 followed individuals on an annual basis up until 2011 and then began collecting information every other year. The 2013 data is the most recently released data from the BLS.

students on an annual basis.

I am able to use the large representative sample of the NLSY97 to compute the probability that a student will complete a four-year college degree by 2013 based on characteristics identified in the previous literature as being predictive of college success. This calculated probability is then used with the reported expectation for each of the students in a subset of the data to create two measures of alignment. I use these measures to examine the following questions. First, are there any observable differences between those who are aligned and misaligned? Next, do respondents become more aligned as they age and progress out of high school? Last, are those who are more aligned at an early age more likely to reach their outcomes?

I find that although the majority of students are overconfident, or more aspirational, in their belief that they will complete a four-year college degree, there are considerable differences in alignment based on several observable characteristics. In most cases the alignment of a student's expectations differ based on their: parent's education, ASVAB percentile, school enrollment, and race; alignment also differs based on other factors depending upon the age of the respondent and when they were asked. Using the two sub-samples of different aged respondents I show that as students age and acquire more information their expectation of completing college becomes more aligned with their estimated probability of completion, but the majority remain overconfident. I confirm this by examining 700 students who are asked their expectations first in 1997 while in high school then again five years later when they either are in college or the workforce. Last I show that those who are more aligned in either direction with what a model of college completion predicts the more likely they are to eventually reach that outcome.

Future research is needed to examine and identify how we can increase student alignment prior to them making the decision of whether or not to attend college. If we can help students more accurately forecast their probability of completing college we

could reduce the number of students who take out burdensome loans and are unable to complete a college degree. It would also be beneficial for schools to identify students who at the threshold are more likely to struggle initially and drop out. Schools could then intervene with proactive measures to prevent this.²³ It also is necessary to examine why there are such distinct differences in the alignment of students based on race and what impact this might have on their future success.

²³The University Innovation Alliance is a program that uses predictive analytics and proactive advising interventions to increase semester-to-semester retention rates and reduce time-to-degree for graduating students. It has just recently been established and currently on includes 11 Universities across the United States.

Chapter 6 A Quantile Analysis of Student Expectations of College Completion and Earnings in their Late 20s to Early 30s

6.1 Introduction

The predictors of future income have been widely studied in the economics literature and a large portion of it focused on human capital and the returns to education.¹ While education has been shown to increase human capital and thus increase earnings, dealing with the issue of selection into higher education has been a problem that constantly plagued researchers and has yet to be completely settled.² The previous essays have shown that a student's expectation of completing college contains some private information that is positively associated with their ability to complete college. It is possible that this expectations data could include some form of person specific information that relates to specific human capital which could be helpful in mitigating the selection issue.

The National Longitudinal Study of 1997 (NLSY97) surveys students first between the ages of 12 and 17 and then follows them annually up until 2011 where-after they are surveyed on a biannual basis. The most recently released data is from 2013 when the respondents were in their late 20s to early 30s. Among other things, the NLSY97 asked different subsets of the respondents to report their subjective belief that they would complete a college degree by the time they were 30 in 1997 and later in 2001. These data combined with the education, employment, and income data reported by

¹See Willis (1986); Angrist and Keueger (1991); Krueger and Ashenfelter (1994); Polachek and Siebert (1993); Becker (1994); Kane and Rouse (1995); Altonji and Pierret (1996); Card (1999); Tyler et al. (2000); Card (2001); Heckman et al. (2008); Polachek et al. (2008)

²Others have looked at the external returns to schooling, which include but are not limited to Lochner and Moretti (2001); Oreopoulos and Petronijevic (2013); Grossman (2006); Blomquist et al. (2014).

the individuals in 2013 allows me to analyze a number of labor related topics.

First I identify noticeable differences in the frequency of reporting, yearly income, and hours worked for those who predict college success and are successful versus those who do not, as well as those who accurately predict that they will not complete college. I find that those who successfully predicted that they would complete college report receiving any income during 2012 at a much higher rate than those who fail to predict success and those who predict they will not complete college. They also report a higher annual income and work nearly 300 more hours per year than both groups. Those who fail to predict success report at a higher rate, report higher income, and work more hours than those who predict failure but the differences are much smaller than those previously mentioned. I then include these expectations in a wage regression and test whether or not they are associated with wages for those in their late 20s and early 30s.³ The results indicate that when individuals report their expectations between the ages of 15 and 17, they are not associated with future earnings. However, when asked between the ages of 17 and 22 the reported expectations are positively associated with future wages. There are considerable differences based on gender, whether they reported at one of the 3 primary heaping points, and which quantile of the wage distribution the respondents were in.

6.2 Data

This research utilizes the full sample and the two subsets that report their educational expectations in 1997 and 2001. There is considerable attrition or lack of response to the income questions so the data utilized differs from that presented above. The summary statistics for each of the samples used who responded to the income questions in the most recent survey are included in table 6.1. The 1997 and

³Only those who report some earnings during the year of 2012 are utilized, so the estimates are for workers only.

2001 subsets have a similar gender and race composition compared to the full sample. Each are nearly 51% male and 49% female, and are approximately 52% white, 25% black, and 23% Hispanic.⁴ The biggest difference among the three samples remains that the 1997 sub-sample is approximately a year and a half older in 2012 than the other two because of how the participants were selected. Only the oldest $\frac{2}{5}$ th of the sample were asked their expectations in 1997 while no age restrictions were present for the 2001 sub-sample. This difference is observable in a number of the other variables that are age dependent including: household size, percentage of the sample married, and potential years of experience.⁵

⁴A means test across the demographic, household, and residential controls reveal that the two expectations samples are not statically different from the full sample.

⁵Calculated as: $Experience = (age\ in\ 2012 - years\ of\ schooling - 6)$

Table 6.1: NLSY97 Income Summary Statistics

Variable	Full Sample	1997 Sub-Sample	2001 Sub-Sample
Income in 2012	\$ 38,574.00	\$ 42,358.20	\$ 38,697.80
Hours Worked in 2012	1794.65	1821.22	1813.99
Avg. Hourly Wage	\$23.26	\$24.92	\$23.09
Log Avg. Hrly Wage	2.86	2.93	2.84
College Expectation 1997		75.04	n/a
College Expectation 2001		n/a	70.57
Representative Sample	0.766	0.759	0.762
Age in 2012	29.87	31.35	29.83
Male	0.513	0.517	0.509
White	0.537	0.538	0.525
Black	0.253	0.25	0.25
HH Size	3.267	3.375	3.227
Married	0.454	0.518	0.439
ASVAB Percentile	48.314	48.227	48.96
-Imputed ASVAB	0.174	0.184	0.165
Years of Schooling	13.987	14.008	14.017
Potential Years of Experience	11.003	12.494	10.927
High School Degree	0.414	0.427	0.42
2 Year College Degree	0.084	0.085	0.078
4 Year College Degree	0.234	0.228	0.243
Masters Degree	0.077	0.085	0.073
Doctoral Degree	0.019	0.017	0.023
Rural	0.164	0.174	0.177
In MSA, Not Central City	0.526	0.537	0.507
In Central City MSA	0.425	0.408	0.432
North East Census	0.156	0.157	0.155
North Central Census	0.223	0.232	0.221
West Census	0.224	0.22	0.235
Agg, For, Fish Industry	0.007	0.005	0.007
Mining Industry	0.006	0.005	0.007
Utilities Industry	0.007	0.008	0.008
Construction Industry	0.058	0.066	0.053
Manufacturing Industry	0.068	0.073	0.053
Wholesale Trade Industry	0.025	0.03	0.017
Retail Trade Industry	0.098	0.093	0.098
Transportation and Warehouse Industry	0.035	0.034	0.038
Communications Industry	0.022	0.022	0.026
Finance, Insurance, and Real Estate Industry	0.065	0.07	0.058
Professional Industry	0.119	0.116	0.117
Education and Health Industry	0.226	0.233	0.235
Entertainment and Hospitality Industry	0.092	0.074	0.095
Other Service Industry	0.043	0.041	0.049
Public Administration Industry	0.041	0.047	0.044
Special ACS Industry	0.085	0.082	0.09
N	5,205	1,887	1,143

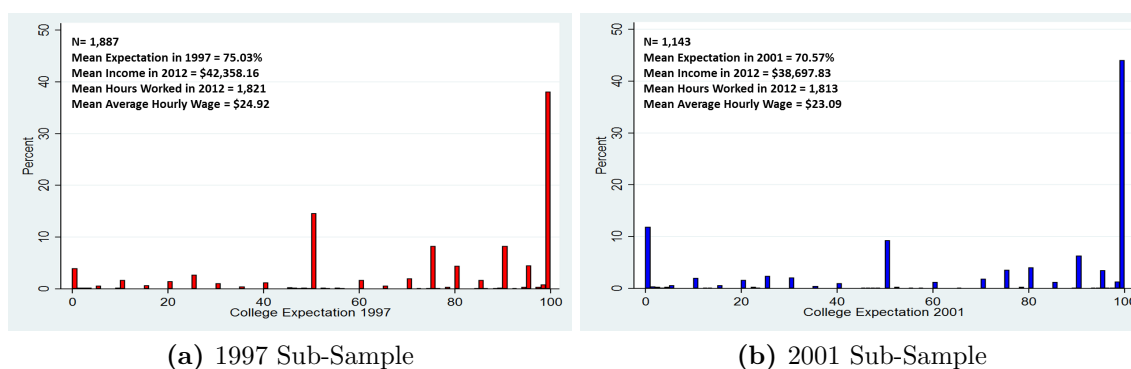
Notes: Only those who reported positive income are utilized in the estimation and presented here. Average hourly wage was calculated as $\frac{\text{Yearly Income}}{\text{Hours Worked}}$

The result of the 1997 sub-sample restricting those who answered the expectations questions is evident when comparing it to the 2001 sub-sample. First, those in the 1997 sub-sample reported higher expectations of completing a college degree by approximately 5 percentage points. The lower expectations for the older cohort is in-line

with the research presented in the essays above.⁶ The 1997 sub-sample also reports a higher average income, number of hours worked, and average hourly wage which are likely the result of their additional potential years of experience in the workforce. I attempt to control for this selection issue by including potential years of experience when estimating whether or not expectations are associated with future earnings in the following sections.

Selection is also evident based on the reporting of income for 2012. For the 1997 sub-sample, 1,887 of the 3,450 respondents reported income in 2012. The mean expectation for these 1,887 was three percentage points higher, 75% vs. 72%, than when the full 1997 sub-sample is used. A similar pattern is present for the 2001 sub-sample with 1,143 of the 1,950 respondents reporting income in 2012, and a 3 percentage point difference, 70% vs. 67%, in the average expectation of completing college. Although those who reported income have a higher mean expectation in both of the sub-samples the distribution of reporting appears similar. The frequency distributions for those who reported income in each sub-sample are included below along with several selected income summary statistics.

Figure 6.1: Distribution of College Completion Expectations and Income Statistics



⁶Hunter (2017a,b,c)

6.3 Analysis of the Wage Data in the NLSY97

The employment and wage data utilized below is from 2013 survey regarding the respondents employment and income during 2012.⁷ The average hourly wage of the respondent is calculated rather than using the reported yearly income in an attempt to deal with issue of unemployment throughout the year and part versus full-time workers.⁸ The data reports whether or not the individual was employed for each of the 52 weeks in the year. I created the number of weeks the individual worked by summing the number of weeks they reported being employed. Overall 34% of the respondents worked for 0 weeks, 47% reported working for 52 weeks, and 19% reported working between 1 and 51 weeks. The 2013 survey also asked those to report the average number of hours they worked each week for up to 10 employers. I used the hours reported from the top three employers to create an average number of hours worked per week for the individual.⁹ If the respondent only reported working for one employer then their average number of hours worked per week was what they reported for that one employer. If they reported working for two employers their average number of hours worked was the sum of their hours working for those two divided by two; a similar process was used for those who reported working for three different employers. Those who reported 0 hours work for all three or working over 100 hours a week on average were not included in the analysis, so the results are for workers only. The average number of hours worked per week for all respondents was 36.5 and ranged from 0 to 96 hours. The average hourly wage of the individual in

⁷Respondents also reported their income for the 2013 calendar year but as participants were interviewed at different dates the data is not comparable without having to estimate their earnings for the remainder of the year.

⁸Future analysis will include data from more years, but the combination of the two year gap between the previous reporting and the young age of respondents maximizes sample size and reliability.

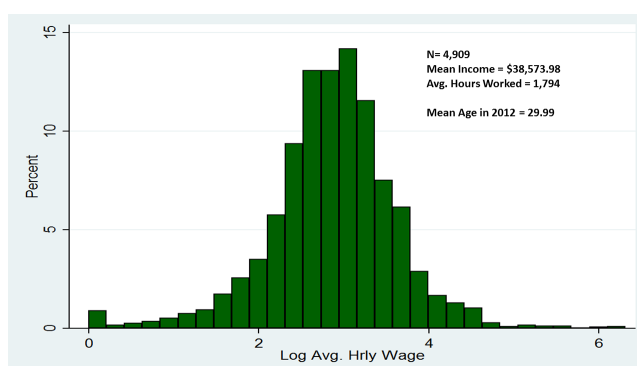
⁹The majority of respondents only reported 1 employer, and only 309 respondents reported working for more than 3 employers.

2012 was then calculated as:

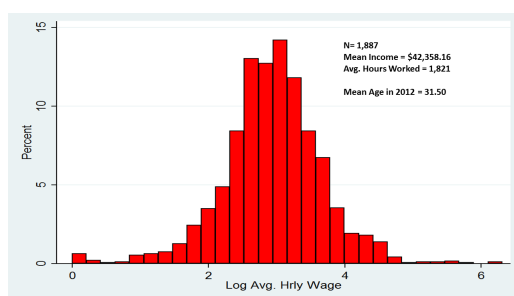
$$Avg. Hourly Wage_{i,2012} = \frac{Total Income_{i,2012}}{(\# Weeks Worked_{i,2012} \times Avg. Hours Worked per Week_{i,2012})}$$

The natural log of this is then taken to reduce the impact of outliers and for later estimating the Mincerian log-lin model.¹⁰ The average wage distribution with some basic summary statistics for the full sample, the 1997 sub-sample, and the 2001 sub-sample are presented below along with a discussion of the work and income variables for each.

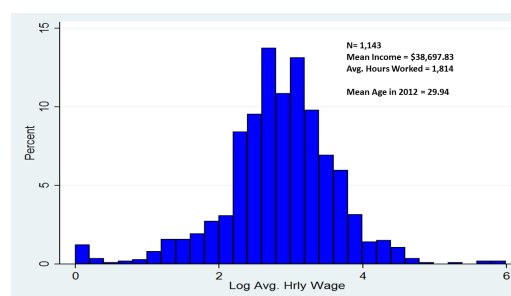
Figure 6.2: Distribution of Ln(Avg. Hourly Wage) in 2012



(a) Full Sample



(b) 1997 Sub-sample



(c) 2001 Sub-sample

Tables 6.2 and 6.3 below include summary statistics of the income and work variables and are separated based on the successfulness of the respondent in predicting

¹⁰Mincer (1974); Polachek et al. (2008)

their future college success. I use these statistics for several reasons. First, I compare the full sample to each of the sub-samples to present the differences caused by the limited sampling in 1997. There are noticeable differences between the full and 1997 sub-sample. On average the two groups report receiving any income at a similar rate of 81% but the 1997 sub-sample reports higher average earnings, number of hours worked, and a higher hourly wage. This is the result of 1997 sub-sample individuals having had more time in the labor market than the younger participants. The full and 2001 sub-sample appear to be more similar in terms of income and work reporting. This suggests that estimates from the models that utilize this sub-sample are more representative than those that use the 1997 sub-sample which should be considered when discussing the findings below.

Next, I examine the mean differences in reporting conditional on how successful the respondent was at predicting whether or not they would complete college. Within each sub-sample I begin by categorizing each student as successfully predicting college completion, unsuccessfully predicting college completion, or successfully predicting that they would not complete college. Those who reported expectations between 75% and 100% and went on to complete a four-year college degree were classified as successfully predicting college completion. Those who reported within the same range but failed to complete a four-year college degree were classified as unsuccessfully predicting college completion. Last, those who reported below a 50% chance of completion and did not go on to complete the degree were classified as successfully predicting that they would not complete college.¹¹

Differences in both the frequency and magnitude of the reported wage statistics are present in both sub-samples across the three prediction categories. Those who successfully predicted that they would complete college report earning any income at

¹¹Robustness checks based on how a successful prediction is classified is still needed.

a much higher rate, have a much larger yearly income, and work many more hours than those who were unsuccessful in predicting college completion and those who predicted they would not complete college. This is not surprising since those with a college degree are much more likely to be employed and are qualified for more high paying jobs.¹² Differences between those who were confident that they would complete college but fail to do so and those who successfully predicted they would not complete college are more intriguing as both groups do not have a college degree so they should be similarly qualified in terms of employment opportunities.¹³ The high expectation unsuccessful predictors report having any income at higher rate, have higher incomes, and work more hours than those who had low expectations and did not complete college. Below I use regression analysis to test whether expectations are associated with early career earnings. However, the anecdotal evidence suggests differences in employment outcomes based on differences in both the magnitude and accuracy of expectations. The differences in expectations might indicate differences in the cost-benefit analysis of students when deciding whether or not to attend college. This will be the subject of future analysis.

¹²Card (1999); Oreopoulos and Petronijevic (2013); Blomquist et al. (2014)

¹³Further analysis that also controls for the number of years of college that each completes is needed.

Table 6.2: Income, Work, and Prediction Summary Statistics for 1997 Sub-sample

Variable	Full Sample	1997 Sub-sample	Successfully Predict College Completion	Unsuccessfully Predict College Completion	Successfully Predict No College Completion
Report Any Income in 2012	81.9% (38.5)	80.8% (39.4)	92.8% (25.9)	77.4% (41.8)	71.8% (45.1)
N	7092	2732	677	1080	436
Yearly Income in 2012	\$38,127.50 (30343.90)	\$41,745.80 (34114.70)	\$56,710.10 (39928.20)	\$36,763.00 (31321.70)	\$32,857.00 (27426.40)
N	5205	1997	583	744	274
Hours Worked in 2012	1622.32 (804.24)	1631.23 (819.67)	1868.09 (760.84)	1573.79 (808.21)	1452.46 (878.61)
N	6140	2374	634	917	360
Hourly Wage in 2012	23.18 (26.52)	24.87 (28.00)	31.83 (24.63)	21.16 (16.85)	20.02 (20.79)
N	4927	1891	565	700	256

Notes: Standard deviation in parenthesis. Those who reported between a 75% and 100% chance of completing a four-year college degree by the time they are 30 are classified as predicting college success, while those who reported lower than 50% are classified as not predicting college success.

Table 6.3: Income, Work, and Prediction Summary Statistics for 2001 Sub-sample

Variable	Full Sample	2001 Sub-sample	Successfully Predict College Completion	Unsuccessfully Predict College Completion	Successfully Predict No College Completion
Report Any Income in 2012	81.9% (38.5)	82% (38.2)	94% (23.6)	82% (38.8)	73% (44.5)
N	7092	1630	439	531	436
Yearly Income in 2012	\$38,127.50 (30343.90)	\$38,362.10 (30845.3)	\$52,721.00 (35040.4)	\$33,174.60 (26344.9)	\$29,226.50 (26184.1)
N	5205	1211	386	391	277
Hours Worked in 2012	1622.32 (804.24)	1642.31 (816.53)	1835.66 (783.76)	1562.32 (772.55)	1537.84 (880.07)
N	6140	1415	414	456	352
Hourly Wage in 2012	23.18 (26.52)	22.991 (24.79)	30.834 (30.23)	20.464 (20.11)	17.166 (12.96)
N	4927	1148	373	362	264

Notes: Standard deviation in parenthesis. Those who reported between a 75% and 100% chance of completing a four-year college degree by the time they are 30 are classified as predicting college success, while those who reported lower than 50% are classified as not predicting college success.

6.4 Estimation and Results

I begin by estimating a simple log-lin Mincerian model. The chart below outlines the model and the covariates used in the estimation. There is concern that the two sub-samples might not be representative of the NLSY97 due to differences in the

selection of those who were asked the expectations questions. To test this I start by estimating several different models and compare estimates across them. First, I use the entire NLSY97 sample and estimate the model excluding the expectations variable and only include the specified covariates. I then estimate a similar model using both the 1997 and 2001 sub-samples first without the expectations included and then with them. Last I estimate the model using the 415 respondents who reported their expectations in 1997 and 2001 and who reported income in 2012.¹⁴

Table 6.4: Earnings Model Description

<u>Model Estimated:</u>	$Ln(Avg.HourlyWage_{i,2012}) = \alpha + \beta Expectation_{i,t} + \theta X_{i,2012} + \epsilon_i$	
<u>Variable of Interest:</u>	The percent chance the student believes that they will obtain a 4 year college degree by the time they are 30	
<u>Control Variables:</u>	Demographics - Male - White - Black	Location - Urban - MSA Central City - MSA Non-Central City - 3 Census Regions (W, NE, NC)
	Ability, Experience, and Household - ASVAB Percentile - Experience and Experience ² - Marital Status	
	Schooling - Years of Schooling - Indicators for Degree Attainment	Industry - 1 of the 16 primary ACS industry codes for the respondent's primary employer

Notes: In the model, the t indicates which sub-sample that is used when estimating. It is not meant to indicate a panel analysis. For each categorical variable, that which is not listed is used as the base case.

The estimates from each of the specified models are included in the table below. Column 1 includes those from the full NLSY97 sample. Columns 2 and 3 contain those from the 1997 sub-sample while columns 4 and 5 include those from the 2001

¹⁴The model estimated includes controls for years of schooling and levels of schooling as well as the ASVAB percentile. These are endogenous and could potentially bias the estimates because of selection and differences in skills that schooling proxies for per (Neal and Johnson, 1996); per (Bollinger, 2003) measurement error in the ability measure could also result in biased estimates. Neal and Johnson (1996) tests a model that includes these controls and are the results do not differ drastically from a simplified model of exogenous characteristics. These will be include in a later analysis.

sub-sample; the first of each are from the specifications that exclude the individual's expectations. The last two columns includes estimates from the 415 who answered the expectations questions in both 1997 and 2001. Across each of the models estimated the male and marital status coefficients are positive and statistically significant; specifications that estimate using men and women separately are discussed in a later section.¹⁵ Excluding the overlapping sub-sample estimates, the ASVAB and schooling controls are positive and statistically significant predictors of wage. Informally testing how representative each of the three samples used are, I compare the estimates from columns 1, 2, and 4 which use each and does not include the expectations variables. In each model the magnitude and statistical significance of the coefficients suggest that both the 1997 and 2001 sub-samples are similar to the full NLSY97 sample. Therefore, the findings from the models that include the expectations variable should be as representative as all of those who report earnings in the NLSY97.

Comparing the specifications that include the individual's expectation versus those that do not, regardless of which sub-sample is used the explanatory power increases when the expectations variable is included. From the estimates it appears that when between 15 and 17 years old the responses do not contain useful information in terms of future hourly wage. When the older 2001 sub-sample is utilized, expectations are positive and significantly associated with future wage. The estimate from column 5 indicates that a 10 percentage point increase in the expectation that one will complete a four-year college degree by the time they turn 30 is associated with a 2.03% increase in their future average hourly wage. Recall that this sub-sample includes respondents who were between 17 and 22 when answering the expectations question and varied in enrollment status. This suggests that students are learning and updating as they age based on their position in life, whether it be in post-secondary education or in

¹⁵The estimates for those married and single separately are not included but can be obtained by emailing the author and will be included in future analysis.

the workforce, and their expectations contain some form of private information that is associated with their future earnings that is captured by the standard economic controls. This notion is reinforced by the estimates from the overlapping sub-sample which includes only the older portion of the sample. They were between 15 and 17 when they initially reported their expectations in 1997 and between 20 and 22 when they reported in 2001. The estimates from column 7 suggest that the respondents early expectations are not predictive of future wage, but after 5 more years of experiencing the world and gathering information their expectations are positive and statistically significant.¹⁶

¹⁶The estimate is similar to the estimate from the full 2001 sub-sample.

Table 6.5: Hourly Wage and Expectations Estimates

$Y = \ln(\text{Avg.HourlyWage}_{i,2012})$							
	Full Sample	1997 Sub-sample		2001 Sub-sample		Overlapping Sub-sample	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
College Expectation 1997			0.000902 (0.000606)				-0.000856 (0.00123)
College Expectation 2001					0.00203** (0.000682)		0.00220* (0.000995)
Male	0.150*** (0.0206)	0.160*** (0.0330)	0.164*** (0.0330)	0.191*** (0.0453)	0.197*** (0.0452)	0.202** (0.0758)	0.208** (0.0756)
White	0.000370 (0.0311)	0.0530 (0.0488)	0.0559 (0.0489)	0.00361 (0.0647)	0.0143 (0.0648)	0.131 (0.0974)	0.147 (0.0984)
Black	-0.114*** (0.0308)	-0.0904 (0.0478)	-0.0943* (0.0478)	-0.0354 (0.0652)	-0.0384 (0.0649)	-0.108 (0.0997)	-0.109 (0.100)
Experience	0.0405** (0.0136)	0.0677 (0.0367)	0.0707 (0.0367)	0.0362 (0.0339)	0.0448 (0.0338)	0.00786 (0.106)	0.0223 (0.105)
Experience ²	0.0000332 (0.000606)	-0.000672 (0.00116)	-0.000717 (0.00116)	0.000207 (0.00158)	0.00000906 (0.00156)	-0.00121 (0.00342)	-0.00168 (0.00338)
Married	0.160*** (0.0204)	0.167*** (0.0320)	0.164*** (0.0322)	0.229*** (0.0452)	0.236*** (0.0450)	0.177* (0.0759)	0.194* (0.0774)
ASVAB Percentile	0.00288*** (0.000469)	0.00313*** (0.000781)	0.00303*** (0.000784)	0.00321** (0.00100)	0.00286** (0.00101)	0.000761 (0.00195)	0.000205 (0.00200)
Years of Schooling	0.0504*** (0.00918)	0.0549* (0.0274)	0.0549* (0.0274)	0.0551** (0.0203)	0.0513* (0.0201)	-0.00210 (0.0610)	-0.00803 (0.0606)
High School Degree	0.148*** (0.0335)	0.174** (0.0535)	0.168** (0.0538)	0.168* (0.0792)	0.152 (0.0787)	0.166 (0.128)	0.176 (0.128)
2 Year College Degree	0.259*** (0.0487)	0.320*** (0.0807)	0.313*** (0.0809)	0.310** (0.109)	0.285** (0.108)	0.194 (0.173)	0.184 (0.171)
4 Year College Degree	0.391*** (0.0499)	0.414*** (0.0833)	0.401*** (0.0836)	0.349** (0.115)	0.301** (0.114)	0.309 (0.216)	0.285 (0.215)
Masters Degree	0.475*** (0.0658)	0.486*** (0.111)	0.476*** (0.111)	0.462** (0.148)	0.434** (0.146)	0.514 (0.289)	0.516 (0.286)
Doctoral Degree	0.629*** (0.0951)	0.882*** (0.162)	0.877*** (0.163)	0.626*** (0.187)	0.602** (0.186)	0.824* (0.332)	0.850* (0.331)
Industry Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4853	1860	1860	1129	1129	415	415
Adjusted R^2	0.218	0.214	0.215	0.231	0.237	0.243	0.246

Notes: Industry controls include which 1 of the 16 primary ACS industry codes the respondent's primary employer was classified as in 2012. Location controls include: whether the respondent lived in a rural or urban area; whether or not they live in an MSA, and if so if it were in the central city or not; and which of the 4 primary census regions the respondents lived in at the time of the survey.

- Robust standard errors in parentheses

- * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Estimates from similar specifications separated by gender for the full, 1997, and 2001 sub-samples are included in the table below.¹⁷ Again the estimates from the

¹⁷This was done in an attempt to deal with gender specific endogenous sorting and has been done in previous research. Per the suggestion of Dr. Jim Ziliak and see Pencavel (1986); Mroz (1987); Lundberg (1988); Neal and Johnson (1996); Lundberg et al. (1997); Cornwell and Rupert (1997); Ginther and Zavodny (2001); Bollinger (2003); Light (2004); Antonovics and Town (2004)

1997 sub-sample suggest that the expectations of the younger cohort are not associated with future wages, but those from the older cohort are. The estimated association at the mean appears to be similar for both male and females.¹⁸ While the mean analysis suggests a positive association between expectations and future hourly wage, further analysis of a possible non-linear impact and differences across the wage distribution could be beneficial. The next section includes a discussion of several non-linear associations tested and results from a quantile regression model.

¹⁸A model with gender interactions was also tested and findings were similar. See the appendix for estimates.

Table 6.6: Hourly Wage and Expectations - Male and Female Estimates

$Y = \ln(Avg.HourlyWage_{i,2012})$						
	<u>Full Sample</u>		<u>1997 Sub-sample</u>		<u>2001 Sub-sample</u>	
	(1-M)	(1-F)	(2-M)	(2-F)	(3-M)	(3-F)
College Expectation 1997			0.00101 (0.000915)	0.000799 (0.000878)		
College Expectation 2001					0.00214* (0.000950)	0.00224* (0.00103)
White	-0.00979 (0.0440)	0.0108 (0.0488)	0.0243 (0.0685)	0.102 (0.0764)	-0.0103 (0.0905)	0.00524 (0.100)
Black	-0.201*** (0.0485)	-0.0545 (0.0441)	-0.193* (0.0784)	-0.0174 (0.0666)	-0.188 (0.108)	0.0601 (0.0811)
Experience	0.0344 (0.0203)	0.0454* (0.0201)	0.0654 (0.0592)	0.0840 (0.0568)	0.112* (0.0455)	-0.0277 (0.0399)
Experience ²	0.000189 (0.000864)	-0.000111 (0.000946)	0.000747 (0.00164)	-0.00242 (0.00198)	-0.00274 (0.00194)	0.00320 (0.00170)
Married	0.211*** (0.0311)	0.111*** (0.0280)	0.216*** (0.0500)	0.115** (0.0443)	0.325*** (0.0717)	0.119 (0.0610)
ASVAB Percentile	0.00179* (0.000699)	0.00438*** (0.000702)	0.00219 (0.00118)	0.00402*** (0.00120)	0.00113 (0.00160)	0.00455** (0.00143)
Years of Schooling	0.0531*** (0.0134)	0.0482*** (0.0135)	0.105* (0.0452)	0.0114 (0.0367)	0.0718* (0.0320)	0.0442 (0.0260)
High School Degree	0.191*** (0.0469)	0.0921 (0.0516)	0.168* (0.0821)	0.128 (0.0719)	0.171 (0.106)	0.0601 (0.117)
2 Year College Degree	0.197** (0.0664)	0.311*** (0.0761)	0.232 (0.125)	0.378*** (0.111)	0.331* (0.159)	0.189 (0.149)
4 Year College Degree	0.365*** (0.0699)	0.404*** (0.0745)	0.299* (0.125)	0.471*** (0.119)	0.329 (0.171)	0.166 (0.160)
Masters Degree	0.309* (0.123)	0.511*** (0.0922)	0.0504 (0.231)	0.672*** (0.150)	0.482* (0.221)	0.229 (0.201)
Doctoral Degree	0.557*** (0.143)	0.655*** (0.130)	0.555* (0.271)	1.053*** (0.200)	0.657* (0.284)	0.448 (0.246)
Industry Controls	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2483	2370	959	901	576	553
Adjusted R^2	0.169	0.244	0.154	0.241	0.221	0.244

Notes: Industry controls include which 1 of the 16 primary ACS industry codes the respondent's primary employer was classified as in 2012. Location controls include: whether the respondent lived in a rural or urban area; whether or not they live in an MSA, and if so if it were in the central city or not; and which of the 4 primary census regions the respondents lived in at the time of the survey.

Marginal effects; Standard errors in parentheses

– * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Non-Linearities and a Quantile Regression Analysis

I test several non-linear relationships between expectations and future wages. First, both a quadratic and cubic relationship between expectations and wage were tested and both were found to be insignificant regardless of which sub-sample was used. Next, the interaction of expectations and various demographic characteristics were tested. The estimates indicate that expectations and hourly wage appear to have a similar association for men and women, as well as whites, blacks, and Hispanics.¹⁹ Indicators for those who reported expectations at the certainty points of 0% and 100% as well as the uncertainty point of 50% were included and found to have varying impacts based on the sub-sample used. Estimates from a quantile analysis that included these indicators also suggested differences across the wage distribution. The results from each of these are discussed after the basic quantile regression model.

The estimates from the quantile regression model, along with those from OLS, are included in the table below.²⁰ For this analysis I examine the .25, .5 (median), and .75 quantiles and compare them to the previous mean estimates from above.²¹ The OLS and quantile estimates from the 1997 and 2001 sub-samples are included in the top portion of the table while the estimates from each sub-sample separated by gender are include in the bottom two portions. While the mean estimates from the 1997 sub-sample are insignificant the quantile analysis suggests a possible association for those in the .75 quantile, especially for women. The full 1997 sub-sample estimates for the upper quantile indicate a 10 percentage point increase in the expectation that the individual will complete college is associated with a 1.48% increase in future hourly

¹⁹This is surprising as previous research has shown that blacks are more overconfident in their ability to complete college than whites and Hispanics. Results from the Neal and Johnson (1996) specifications also suggest possible variations from this, including for the 1997 sub-sample.

²⁰The controls used above included in each of the models but are excluded from the charts for efficiency in presentation. The full estimates can be found in the appendix.

²¹The sample size, especially when separated by gender, limits a more detailed analysis.

wage. When the model is estimated separately for men and women the estimates suggest between a 1% and 2% increase in future hourly wage, but only for women in the median and upper quantile.

Estimates across each quantile from the full 2001 subs-sample suggest a similar association as those from the OLS estimates. However, when the model is estimated separately for men and women the estimates vary considerably. First, for women, expectations appear to have the most impact at the mean and the upper quantile. These estimates indicate that a 10 percentage point increase in a woman's expectation is associated with around a 2% increase in their future hourly wage. Expectations for the lowest quantile and the median are not statistically significant. For men, only the estimates at the mean and the lowest quantile are statistically significant and have a similar magnitude as those from the female sample. The estimates indicate that expectations of completing college are associated with future early adult-hood income, but there is considerable variation based on gender and quantile. Further analysis is needed to discover why these differences are present.

Table 6.7: Quantile and OLS Estimates of Hourly Wage and Expectations

$Y = \ln(\text{Avg.HourlyWage}_{i,2012})$								
	<u>1997 Sub-sample</u>				<u>2001 Sub-sample</u>			
	OLS	$\tau = .25$	$\tau = .5$ (Median)	$\tau = .75$	OLS	$\tau = .25$	$\tau = .5$ (Median)	$\tau = .75$
College Expectation 1997	0.000912 (0.000623)	0.000641 (0.000470)	0.000486 (0.000432)	0.00148** (0.000541)				
College Expectation 2001					0.00201** (0.000691)	0.00156** (0.000533)	0.00191*** (0.000554)	0.00188*** (0.000545)
Standard Controls Included	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1860	1860	1860	1860	1129	1129	1129	1129
Adjusted R^2	0.193				0.221			
Pseudo R^2		0.1461	0.1437	0.1379		0.1719	0.1713	0.1595

	<u>Male Respondents in 1997 Sub-sample</u>				<u>Female Respondents in 1997 Sub-sample</u>			
	OLS	$\tau = .25$	$\tau = .5$ (Median)	$\tau = .75$	OLS	$\tau = .25$	$\tau = .5$ (Median)	$\tau = .75$
College Expectation 1997	0.00101 (0.000915)	0.00106 (0.000757)	-0.000166 (0.000612)	-0.00132 (0.000877)	0.000799 (0.000878)	-0.000501 (0.000656)	0.00111* (0.000548)	0.00198** (0.000752)
Standard Controls Included	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	959	959	959	959	901	901	901	901
Adjusted R^2	0.154				0.241			
Pseudo R^2		0.137	0.131	0.111		0.179	0.179	0.183

	<u>Male Respondents in 2001 Sub-Sample</u>				<u>Female Respondents in 2001 Sub-Sample</u>			
	OLS	$\tau = .25$	$\tau = .5$ (Median)	$\tau = .75$	OLS	$\tau = .25$	$\tau = .5$ (Median)	$\tau = .75$
College Expectation 2001	0.00214* (0.000950)	0.00126 (0.000786)	0.00146 (0.000775)	0.00262** (0.000934)	0.00224* (0.00103)	0.00245** (0.000851)	0.00116 (0.000669)	0.00183 (0.00122)
Standard Controls Included	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	576	576	576	576	553	553	553	553
Adjusted R^2	0.221				0.244			
Pseudo R^2		0.2033	0.1766	0.1525		0.2030	0.1960	0.1929

Notes: The standard economic controls previously presented are all included in each of the regression models but are excluded from the table for and increase in presentability of the table. See the Appendix for the full results.

Marginal effects; Robust standard errors in parentheses

– * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The last table includes estimates from the specifications that include indicators for those who reported 0%, 50%, or 100%. There is considerable heaping at these three values with the largest percentage of respondents reporting at 100% in both sub-samples. There is the possibility that those who report at any of these points are

choosing them as a way of avoiding the question.²² Also, those who choose to report at these values could be noticeably different than those who report at others as they might be thinking more in-depth about the question or have a better understanding of probability. The table is laid out similarly to the one above with the OLS and quantile estimates from the 1997 and 2001 sub-samples included in the top portion of the table while the estimates from each sub-sample separated by gender are included in the bottom portions.

Similar to the previous estimates, when using the full 1997 sub-sample expectations appear to be unrelated to future hourly wage. When the full 2001 sub-sample is used the continuous measure of expectations is no longer significant across all of the specifications exuding the .25 quantile. However, contrary to the above hypothesis it appears that those who report a 100% chance and are in either the median or upper portion of the wage distribution earn nearly 9% more than those who reported a 99% chance. This could be that those who reported a 100% chance and who were successful in predicting their completion earned higher wages than those who reported a 100% chance and did not complete the degree.

When the models are estimated separately for men and women the estimates change drastically. For the 1997 sub-sample, the men who reported a 100% chance and were not in the .75 quantile had at least 14% higher earnings than those who reported a 99% chance; the largest impact was for those in the .25 quantile with a 24.9% increase. The continuous measure nor any of the indicators were significant for the women in this sub-sample. The only significant estimates from the men in 2001 sub-sample were for those who reported either a 0% or 100% chance and who were in the .75 quantile. Those who reported a 0% chance here had an hourly wage 20% lower than those who reported a 1% chance of completing college. Those who reported a

²²My previous research suggest this to be the case for a large portion of those who report a 100% chance.

100% chance had a 17% higher average hourly wage than those who reported a 99% chance. For the women of the 2001 sub-sample, only those who were in the .25 or .5 quantile and reported a 100% chance of completing college had statistically different hourly earnings. Their wages were between 14% and 19% higher than those who also were in a similar quantile and reported a 99% chance of completing college.

Table 6.8: Quantile and OLS Estimates of Hourly Wage and Expectations with Indicators

$Y = \ln(\text{Avg.HourlyWage}_{i,2012})$								
	1997 Sub-sample				2001 Sub-sample			
	OLS	$\tau = .25$	$\tau = .5$ (Median)	$\tau = .75$	OLS	$\tau = .25$	$\tau = .5$ (Median)	$\tau = .75$
College Expectation 1997	0.000239 (0.000868)	-0.000833 (0.000622)	-0.000450 (0.000642)	0.000964 (0.000635)				
College Expectation 2001					0.00165 (0.00119)	0.00197* (0.000909)	0.000482 (0.000922)	0.00126 (0.000914)
Expectation =0	-0.00715 (0.116)	-0.0727 (0.0554)	-0.106 (0.0719)	0.255 (0.196)	0.0601 (0.105)	0.101 (0.0810)	-0.0630 (0.0868)	-0.0111 (0.0894)
Expectation = 50	0.00252 (0.0513)	-0.0392 (0.0373)	0.0151 (0.0379)	0.00441 (0.0326)	-0.0232 (0.109)	0.0622 (0.0786)	0.00690 (0.0489)	-0.0487 (0.0827)
Expectation = 100	0.0661 (0.0434)	0.112** (0.0347)	0.0517 (0.0364)	0.0572 (0.0404)	0.0941 (0.0591)	0.0414 (0.0486)	0.0895* (0.0410)	0.0887* (0.0431)
Standard Controls Included	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1860	1860	1860	1860	1129	1129	1129	1129
Adjusted R^2	0.193				0.221			
Pseudo R^2		0.149	0.145	.140		0.173	0.173	0.161

	Male Respondents in 1997 Sub-sample				Female Respondents in 1997 Sub-sample			
	OLS	$\tau = .25$	$\tau = .5$ (Median)	$\tau = .75$	OLS	$\tau = .25$	$\tau = .5$ (Median)	$\tau = .75$
College Expectation 1997	0.000246 (0.00120)	-0.00189* (0.000897)	-0.00177* (0.000785)	-0.000136 (0.00116)	0.000243 (0.00126)	-0.000773 (0.000966)	0.000997 (0.000848)	0.00146 (0.00115)
Expectation =0	0.119 (0.153)	-0.117 (0.0994)	-0.0460 (0.165)	0.331*** (0.0965)	-0.230 (0.180)	-0.134 (0.120)	-0.134 (0.102)	-0.286 (0.232)
Expectation = 50	0.0299 (0.0706)	-0.00456 (0.0649)	0.0311 (0.0467)	0.0549 (0.0573)	-0.0326 (0.0698)	-0.0426 (0.0573)	0.00636 (0.0464)	-0.0827 (0.0564)
Expectation = 100	0.154* (0.0693)	0.249*** (0.0401)	0.145*** (0.0450)	0.129 (0.0889)	-0.0236 (0.0551)	-0.0293 (0.0479)	-0.0323 (0.0480)	-0.0244 (0.0388)
Standard Controls Included	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	959	959	959	959	901	901	901	901
Adjusted R^2	0.158				0.241			
Pseudo R^2		0.147	0.135	0.120		0.181	0.181	0.185

	Male Respondents in 2001 Sub-sample				Female Respondents in 2001 Sub-sample			
	OLS	$\tau = .25$	$\tau = .5$ (Median)	$\tau = .75$	OLS	$\tau = .25$	$\tau = .5$ (Median)	$\tau = .75$
College Expectation 2001	0.00171 (0.00144)	0.000694 (0.00186)	0.00108 (0.000943)	-0.000207 (0.00133)	0.00270 (0.00194)	0.00222 (0.00155)	0.000699 (0.00117)	0.00183 (0.00157)
Expectation =0	-0.0133 (0.130)	-0.0632 (0.156)	-0.126 (0.0807)	-0.210* (0.101)	0.208 (0.172)	0.0355 (0.150)	0.0622 (0.115)	0.264 (0.185)
Expectation = 50	-0.131 (0.158)	-0.0132 (0.116)	-0.0938 (0.112)	-0.0145 (0.0965)	0.160 (0.124)	0.155 (0.0989)	0.128 (0.0806)	0.0312 (0.0911)
Expectation = 100	0.0496 (0.0854)	0.0220 (0.0804)	-0.0418 (0.0606)	0.177* (0.0697)	0.127 (0.0838)	0.147** (0.0459)	0.189*** (0.0486)	0.0750 (0.0664)
Standard Controls Included	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	576	576	576	576	553	553	553	553
Adjusted R^2	0.219				0.248			
Pseudo R^2		0.203	0.178	0.157		0.208	0.206	0.197

Notes: The standard economic controls previously presented are all included in each of the regression models but are excluded from the table for and increase in presentability of the table. See the Appendix for the full results.

Marginal effects; Robust standard errors in parentheses

– * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

6.5 Conclusions

The predictors of future income have been widely studied in the economics literature and a large portion of it focused on human capital and the returns to education.²³ While education has been shown to increase human capital and thus increase earnings, dealing with the issue of selection into higher education has been a problem that constantly plagued researchers and has yet to be completely settled. The previous essays have shown that a student's expectation of completing college contains some private information that is positively associated with their ability to complete college. It is possible that this expectations data could include some form of person specific information that relates to specific human capital which could be helpful in mitigating the selection issue.

The National Longitudinal Study of 1997 (NLSY97) surveys students first between the ages of 12 and 17 and then follows them annually up until 2011 where-after they are surveyed on a biannual basis. The most recently released data is from 2013 when the respondents were between their late 20s to early 30s. Among other things, the NLSY97 asked different subsets of the respondents to report their subjective belief that they would complete a college degree by the time they were 30 in 1997 and later in 2001. These data combined with the education, employment, and income data reported by the individuals in 2013 allows me to analyze a number of labor related topics.

First I identify noticeable differences in the frequency of reporting, yearly income, and hours worked for those who predict college success and are successful versus those who do not, as well as those who accurately predict that they will not complete college. I find that those who successfully predicted that they would complete college

²³See Willis (1986); Angrist and Keueger (1991); Krueger and Ashenfelter (1994); Polachek and Siebert (1993); Becker (1994); Card (1999, 2001); Polachek et al. (2008)

report receiving any income during 2012 at a much higher rate than those who fail to predict success and those who predict they will not complete college. They also report a higher annual income and work nearly 300 more hours per year than both groups. Those who fail to predict success report at a higher rate, report higher income, and work more hours than those who predict failure but the differences are much smaller than those previously mentioned. I then include these expectations in a wage regression and test whether or not they are associated with wages for those in their late 20s and early 30s.²⁴ The results indicate that when individuals report their expectations between the ages of 15 and 17, they are not associated with future earnings. However, when asked between the ages of 17 and 22 the reported expectations are positively associated with future wages. There are considerable differences based on gender, whether they reported at one of the 3 primary heaping points, and which quantile of the wage distribution the respondents were in.

Further research is needed to analyze the differences in the association of expectations and future hourly wage conditional on the characteristics noted above. This study will also benefit greatly when the 2015 data is released as a reliable second year of income data can be added.²⁵ Directly relating the expectations data to selection into college is also a future avenue of research that could be used to improve the returns to education estimates.

²⁴Only those who report some earnings during the year of 2012 are utilized, so the estimates are for workers only.

²⁵The 2011 data is not utilized as it is income from 2010. A large percentage of respondents were still enrolled in college and there were lingering effects of the financial crisis that could potentially skew the data.

Chapter 7 Conclusion to the Dissertation

This dissertation contributes in several ways to the economics literature by both emphasizing the importance and usefulness of subjective probabilistic data and examining other previously unexplored factors associated with student success in attending and completing college. Although subjective probabilistic data has been utilized in other fields of economics it has not been used in the economics of education realm. Essay one has two primary objectives. The first is to analyze the expectations data reported by high school students to show that individuals as young as fifteen can and do answer subjective probabilistic questions regarding their education in a logical manner. Next, it shows that these data contain private information about the respondent that can be used to predict future success in post-secondary education. To examine the information contained in these expectations I test how they are formed and updated in essay 2; I also test whether or not students update in a Bayesian fashion. Essay three then builds on this by utilizing a model of college completion and examining the accuracy and alignment of the students' expectations. The last essay studies the differences in the earnings distributions based on the reporting and accuracy of the respondents predictions and later shows that the private information in the expectations variable has some predictive power in future hourly wages.

There are several sources of motivation behind this research. First, the rising cost of college, the resulting increase in the level of student debt, and reports of rising unemployment or under-employment by recent college graduates has people reassessing

whether or not college is worthwhile investment.¹ However, what is rarely discussed is the matriculation decision process and elements which contribute to both students' attendance and success at post-secondary institutions. This is even more surprising considering the 6-year graduation rate for first-time, full-time undergraduate students who began seeking a bachelor's degree at a four-year degree-granting institution in the fall of 2009 was 59 percent.²

The second is that obtaining reliable estimates of the returns to education requires that researchers successfully deal with the issues of selection. A popular approach is to model school choice using observed outcome data, assume that students have rational expectations, and that they condition their beliefs on similar variables and process information in a similar way (Manski, 1993). However, students make schooling choices under uncertainty - uncertainty about personal tastes, individual abilities, and realizations of choice-related outcomes (Zafar, 2011). While the decision process has been examined theoretically (Manski, 1989; Altonji, 1993; Hilmer, 1998; Malumud, 2005) and empirically (Bamberger, 1987; Arcidiacono, 2004; Strange, 2012), most rely on these unrealistic assumptions regarding student expectations. Either implicitly or explicitly, researchers assume that individuals are rational and the formation process is homogeneous. However, if this rational expectations assumption is violated then a number of issues arise. My research uses interpretable subjective data on expectations, as called for by Manski (1993), to show that this rational expectations assumption is flawed.

This research takes advantage of a unique set of questions from the 1997 National

¹While increases in the published cost of college has largely been offset by additional grants or other forms of aid, the out of pocket cost of attending college has increased substantially over the past decade. In 2016-17, the estimated average net tuition and fee price paid by full-time in-state students at public four-year institutions was \$3,770 higher than the net price a decade earlier and \$1,550 higher than the low of \$2,200 in 2009-10. - CollegeBoard (2016); Kroeger et al. (2016); Allen and Seaman (2010); Vedder et al. (2013); Cutler and Lleras-Muney (2006)

²Greenstone and Looney (2012); Department of Education (2013); Shapiro et al. (2015)

Longitudinal Survey of Youth (NLSY97). The survey is nationally representative and collects data on individuals annually from 1997 to 2011, and biannually from 2011 to present. Unlike previous studies, I observe students first while they are in high school and then follow them on an annual basis up until 2013 when they are between the ages of 28 and 33.³ In 1997 and 2001, the survey asked different subsets of students to report their subjective probability of obtaining a four-year college degree by the time they turned 30. In 1997 respondents were in high school and in 2001 most had entered either college or the workforce. To date no other studies have used these data to examine how these expectations relate to post-secondary or employment outcomes. The probabilistic responses also improve upon past research as the measures of educational expectations that have been used are less precise. These studies have used data where students either reported the highest level of education they expected to complete without any notion of how confident they were in realizing that outcome or they reported the likelihood of completing different levels of education using a Likert scale.

The first essay presents new estimates of the factors that predict a student's decision to attend college and whether or not they complete a four-year degree that economists have not previously explored. Manski (1993) outlined the importance of understanding and identifying how a student's expectations influence their decision to attend college or not. He notes that the estimated returns to education could be flawed when economists rely on the assumption of rational expectations instead of attempting to understand the process by which students use information to form expectations about the costs and benefits of a college degree. Dominitz and Manski (1996) and Manski (2004) both examine these issues, but are limited by small sample

³The NLSY97 followed individuals on an annual basis up until 2011 and then began collecting information every other year. The 2013 data is the most recently released data from the BLS and is utilized in each of the essays below.

sizes and cross-sectional data. Jacob and Wilder (2011) are able to observe students over time and study the role of expectations in degree attainment, but do not examine whether these also influence the individual's decision to attend college. They are also limited by the nature of their data as their measure of expectations is imprecise. This essay adds to the literature by utilizing subjective probabilistic data for students that are followed as they transition from high school into either college or the workforce and observes whether or not they fulfilled their expectations.

The results indicate that, even after controlling for the factors previously identified in the economics of education literature, a student's self-reported probability of completing a four-year college degree plays a role in predicting whether they attended college and if they were successful in obtaining a four-year degree. Next, as students age and progressed through more schooling, their expectations played a larger role in predicting completion. Parent expectations are also estimated to be influential in these models when students are young, but the impact diminishes as the student ages and gathers information related to the costs and benefits of a college degree. The estimates signal that students' expectations contain a form of private information that is helpful in determining post-secondary schooling attendance and success that isn't being captured by the previously identified controls. These findings suggests that youth and econometricians may possess and utilize different data when analyzing schooling decisions.

The second essay begins by examining whether or not students incorporate the factors that have been identified with college success when forming their expectations. I estimate which factors students utilize when forming their educational expectations using two different aged subsets of the data. I find that while high school aged students take into account several of the relevant factors, they also appear to be neglecting the impact that income and ability have on their likelihood of completing college. Overall, students are extremely overconfident in their ability to complete a four-year

college degree and it is possible that they either are not incorporating these factors into their expectations. I then test whether or not students update their expectations as a Bayesian model predicts. A Bayesian model is presented and I discuss the three ways in which students should respond to the acquisition of new information. I then present four new sources of information that become available between 1997 and 2001 to a subset of respondents who reported their expectations in both periods. I show that while students do appear to be utilizing the new information when updating their expectations, they do not appear to be true Bayesian updaters. I hypothesize two possibilities as to why this is not the case.

Essay 3 examines the accuracy of the students' predictions that they will complete a four-year college degree. The goal is to identify if the accuracy of predictions vary based on observable characteristics and if so how this information can be used to help students form their expectations more accurately. This could serve as a way to increase the graduation rates at post-secondary universities and reduce the percentage of those who take on excessive student loans and drop out. I am able to use the large representative sample of the NLSY97 to compute the probability that each student will complete a four-year degree by 2013 based on characteristics identified in the previous literature. Two new measures of alignment are then created that use this predicted probability and the student's expectation. I use these measures to test whether alignment differs based on observables, improves as the respondent ages, and if it is associated with their realized outcomes. I find that on average, student alignment differs based on age, parent education, ASVAB percentile, school enrollment, and race. As they age and acquire more information, their expectation of completing college becomes more aligned with their estimated probability of completion. Also, those whose expectations were similar to their predicted probability of completing college were more likely to eventually reach their expected outcome.

The last essay includes these educational expectations in a wage regression to test

whether the private information contained within them has some predictive power for future hourly earnings. I begin by identifying differences in the frequency of reporting, yearly income, and hours worked for those who predict college success and were successful versus those who were not, as well as those who accurately predict that they would not complete college. I find that those who successfully predicted that they would complete college reported receiving any income during 2012 at a much higher rate than those who failed to predict success and those who predicted they would not complete college. They also reported a higher annual income and worked nearly 300 more hours per year than both groups. Those who failed to predict completing college reported receiving any income at a higher rate, a higher income, and that they worked more hours than those who predicted they wouldn't complete college, but the differences were marginal. The results from the wage regression indicate that when individuals report these expectations between the ages of 15 and 17, they are not associated with future earnings. However, when asked between the ages of 17 and 22 the reported expectations are positively associated with future wages. There are considerable differences based on gender, whether they reported at one of the three primary heaping points, and which quantile of the wage distribution they were located.

Within each essay I outlined several topics of future research I plan to pursue, but I also believe the addition of the 2015 data when released will also benefit each. The additional year of data will allow for a more detailed analysis of students fulfilling their expectations. The chapter on alignment will also improve as the non-traditional students who take longer to graduate might have gone on to complete college. The final essay would benefit the most as more students should be transitioning to the workforce, more complete earnings data can be included, and the sample-size should increase.

One common critique of this research is that the acquisition of more or "better"

data could result in expectations being insignificant in predicting college success and completion. The argument is that if we are able to collect better data and include it then it might not be that students have private information, but rather economists have incomplete information. While I understand this argument I'm skeptical of whether or not better data would drive the estimated impact of the expectations to zero. First, the influence of the student's expectations persists even with a large number of controls, all of which have been shown to be related to college success, included in the models. It is possible that measures of dedication, grit, persistence, etc.... could also be include and result in a diminished impact of expectations, but I would argue that there would remain some individual-specific component contained within the expectations that still would not be completely captured. It could be how students are assessing the question, in what way they are forming their expectation in terms of what they each believe will or will not make them successful if they attempt to complete a college degree, or some combination of these and other factors. Future research should examine these possibilities, but as it currently stands the data we have now is what we must rely on.

While the findings from above do not prove a causal relationship between expectations and achievement, it does suggest a robust positive relationship. For policy, I would suggest that schools use these findings when helping students plan for their future. They could in several ways incorporate the students' expectations or aspirations into either their planning or recommendations. I would also suggest an in-depth study of why student expectations of completing at least some college have risen so drastically over the past 20+ years while completion rates have remained relatively unchanged. Are we promoting a four-year college degree to some students who would be better served by a two-year or technical degree? Or, are we sending students to a four-year, two-year, or technical school that might be better of joining the workforce directly? I believe both of these, and other related, questions could help directly

improve student outcomes by increasing completion rates. This then indirectly could benefit society as whole due to the spill-over effects associated with a more educated population.

Appendix A Dissertation Appendix

Chapter 1 Appendix

Figure A.1: Summary of the Subsamples in the NLSY97 Utilized

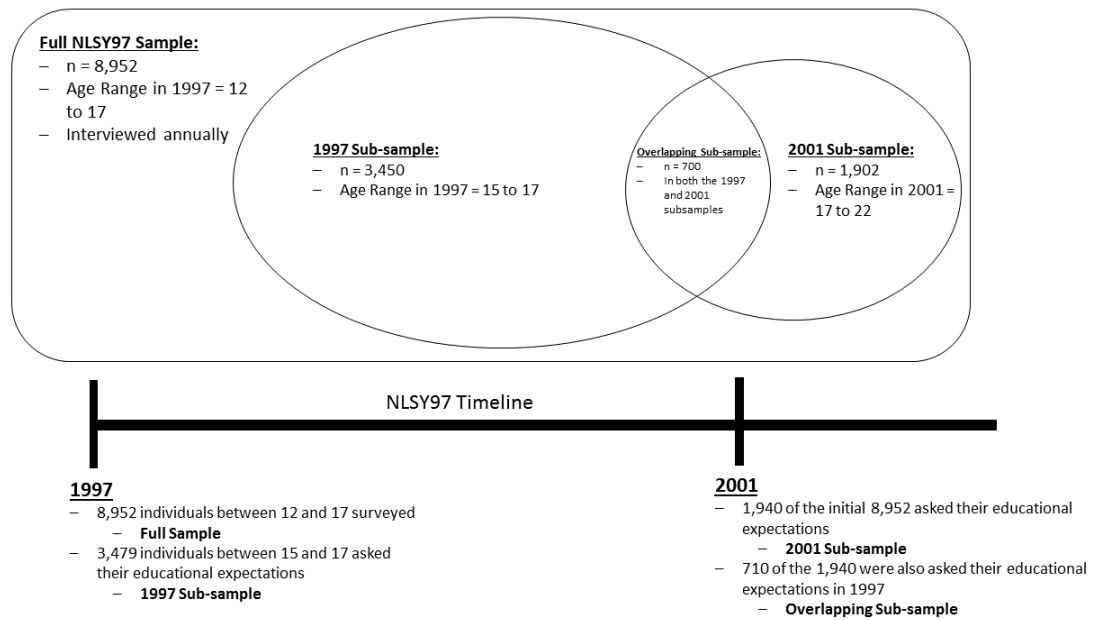


Table A.1: NLSY97 Full and Sub-sample Summary Statistics

	Sample Type	Individuals	Mean Age	Male	White	Black	Hispanic
Full Sample	Cross-Section	6,614	13.92	51.39%	69.44%	16.84%	13.71%
	Over-Sample	2,162	14.02	51.25%	0%	55.64%	44.36%
	Total	8,776	13.98	51.36%	52.34%	26.40%	21.26%
1997 Sub-sample	Cross-Section	2,582	15.47	50.39%	69.71%	16.77%	13.52%
	Over-Sample	868	15.47	51.84%	0%	57.72%	42.28%
	Total	3,450	15.47	50.75%	52.17%	27.08%	20.75%
2001 Sub-sample	Cross-Section	1,440	18.90	51.67%	68.13%	15.90%	15.97%
	Over-Sample	462	18.96	49.13%	0%	55.19%	44.81%
	Total	1,902	18.92	51.05%	51.58%	25.44%	22.98%
Overlapping 700	Cross-Section	517	15.52	49.32%	68.86%	14.89%	16.25%
	Over-Sample	183	15.42	49.18%	0%	55.19%	44.81%
	Total	700	15.49	49.29%	50.86%	25.43%	23.71%

Notes: Mean age for the full sample, the 1997 sub-sample, and the overlapping 700 is as of December 31st, 1996. The overlapping 700 are the 700 respondents who answered the expectations questions in both 1997 and 2001.

Figure A.2: Frequency Distribution of College Completion Expectations

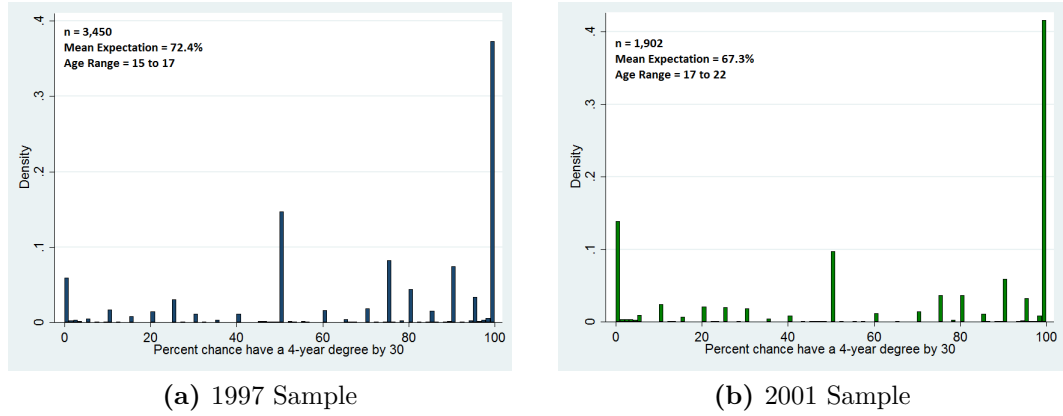


Table A.2: Summary of College Completion Expectations by Sub-sample

	N	Mean	Median	Standard Deviation	Minimum	Maximum
1997	3,450	74.0%	90%	31.2	0%	100%
2001	1,902	65.4%	90%	39.8	0%	100%
1997-2001	700	-8.6	0	40.1	-100	100

Notes: Summary of responses to the following question asked of each sub-sample: "Now think ahead to when you turn 30 years old. What is the percent chance that you will have a four-year college degree by the time you turn 30?" 1997-2001 represents the change in expectations for the 700 individuals that reported expectations in both 1997 and 2001

Table A.3: Student Expectations and Outcomes

Sample	Student Outcomes in 2013			
	Attend College	Does Not Attend College	Completes College	Does Not Complete College
1997 Expectations	84.30% (1,850)	58.53% (1,600)	90.49% (877)	66.17% (2,573)
2001 Expectations				
- \leq 11th Grade	85.59% (333)	41.47% (459)	93.01% (144)	53.38% (532)
- = 12th Grade	83.49% (292)	43.14% (360)	95.95% (120)	52.69% (648)
- $>$ 12th Grade	88.36% (461)	n/a	97.79% (258)	76.93% (200)

Notes: Mean expectations are reported for those associated with each outcome. The number of students in each category is in parenthesis. r below represents the calculated correlation coefficients.

- $r_{attend,exp_97} = .4028$

- $r_{complete,exp_97} = .3320$, $r_{complete,exp_01} = .4587$

Table A.4: Magnitude of Updated Expectations from 1997 to 2001

Increase in Chance Have a 4 Year Degree by 30			Decrease in Chance Have a 4 Year Degree by 30		
	Number of Individuals	Mean 2001 Expectation		Number of Individuals	Mean 2001 Expectation
1 - 25 %	124	89.5%	1 - 25 %	108	52.3%
26 - 50 %	59	85.7%	26 - 50 %	69	27.9%
51 - 75 %	14	93.6%	51 - 75 %	43	12.7%
76 - 90 %	7	94.3%	76 - 90 %	19	7.1%
91 - 99 %	0	N/A	91 - 99 %	9	1.9%
100 %	7	100%	100 %	24	0%
Total	212	89.2%		272	30.4%

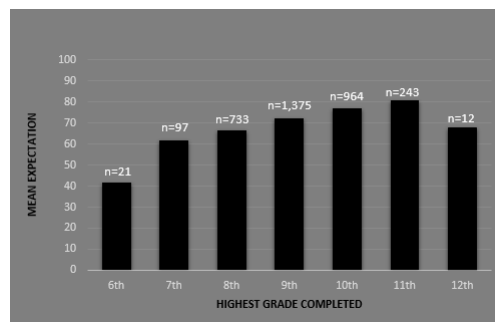
Notes: 216 individuals did not update their expectations from 1997 to 2001, for these individuals the Mean is 86.1% with a Standard Deviation equal to 29.1 percentage points.

Table A.5: College Attendance, Completion, and Expectations

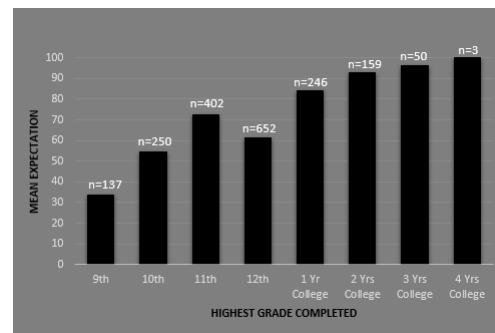
	Percent Chance Have a 4 Year Degree by 30	Number of Students	Number that Attempt Degree	Number that Complete Degree	Percent that Attempt Degree	Percent that Complete Degree	Percent that Start and Complete Degree
1997 Sub-sample	0%	205	37	9	18%	4.4%	24.3%
	1 - 25 %	278	58	5	20.9%	1.8%	8.6%
	26 - 49 %	103	34	3	33%	2.9%	8.8%
	50 %	508	170	37	33.5%	7.3%	21.8%
	51 - 75 %	433	243	92	56.1%	21.2%	37.9%
	76 - 90 %	476	317	149	66.6%	31.3%	47.0%
	91 - 99 %	226	174	112	77.0%	49.6%	64.4%
	100 %	1,221	852	470	69.8%	38.5%	55.2%
	Total	3,450	1,885	877	54.6%	25.4%	46.5%
2001 Sub-sample	0 %	263	39	3	14.8%	1.1%	7.7%
	1 - 25 %	179	41	1	22.9%	0.6%	2.4%
	26 - 49 %	64	12	1	18.8%	1.6%	8.3%
	50 %	184	54	6	29.3%	3.3%	11.1%
	51 - 75 %	126	70	18	55.6%	14.3%	25.7%
	76 - 90 %	211	126	48	59.7%	22.7%	38.1%
	91 - 99 %	143	119	85	83.2%	59.4%	71.4%
	100 %	732	628	360	85.8%	49.2%	57.3%
	Total	1,902	1,089	522	57.3%	27.4%	47.9%

Notes: Number that Attempt Degree includes those who reported their highest grade completed as 1 year of college or more.

Figure A.3: College Completion Expectations by Highest Grade Completed



(a) 1997 Sample



(b) 2001 sample

Table A.6: College Completion Expectations and Attainment by Race and Gender

		N	Mean	Median	Std. Deviation	Percent that Complete College
1997	Male	1,751	68.14%	75%	32.4	20.6%
	Female	1,699	76.69%	94%	30.72	28.5%
	White	1,800	74.42%	90%	31.40	32.3%
	Black	934	73.42%	90%	32.24	18.0%
	Hispanic	716	66.45%	75%	32.11	13.3%
	White Male	918	68.72%	80%	32.54	28.1%
	Black Male	452	70.85%	80%	32.28	14.6%
	Hispanic Male	381	63.54%	70%	32.23	9.5%
	White Female	882	79.80%	95%	29.13	36.7%
	Black Female	482	75.83%	95%	32.06	21.2%
	Hispanic Female	335	69.75%	75%	31.70	17.3%
	Total	3,450	72.35%	85%	31.91	24.5%
2001	Male	971	63.16%	80%	39.08	22.2%
	Female	931	71.68%	95%	37.45	30.3%
	White	981	69.92%	90%	38.15	37.1%
	Black	484	67.24%	90%	38.27	16.5%
	Hispanic	437	61.62%	75%	39.10	12.4%
	White Male	510	65.69%	85%	38.60	31.8%
	Black Male	240	62.83%	80%	39.70	12.5%
	Hispanic Male	221	57.69%	60%	39.12	10.9%
	White Female	471	74.51%	99%	37.15	42.9%
	Black Female	244	71.57%	95%	36.37	20.5%
	Hispanic Female	216	65.64%	85%	38.76	13.9%
	Total	1,902	67.33%	90%	38.52	26.18%

Notes: Expectations are reported between 0 and 100%, so the standard deviation are percentage point changes rather than percent.

Table A.7: College Completion Statistics by Race and Gender

		Number of Students	Number that Attempt Degree	Number that Complete Degree	Percent that Complete Degree	Percent that Complete Degree Conditional on 1 Year Completion
1997	Male	1,751	816	360	20.6%	44.1%
	Female	1,699	1,018	484	28.5%	47.5%
	White	1,800	1,075	582	32.3%	54.1%
	Black	934	441	168	18.0%	38.1%
	Hispanic	716	318	94	13.3%	29.6%
	White Male	918	498	258	28.1%	51.8%
	Black Male	452	171	66	14.6%	38.6%
	Hispanic Male	381	147	36	9.4%	24.5%
	White Female	882	577	324	36.7%	56.2%
	Black Female	482	270	102	21.2%	37.8%
	Hispanic Female	335	171	58	17.3%	33.9%
	Total	3,450	1,834	844	24.5%	46.0%
2001	Male	971	492	216	22.2%	43.9%
	Female	931	581	282	30.3%	48.5%
	White	981	627	364	37.1%	58.1%
	Black	484	252	80	16.5%	31.7%
	Hispanic	437	194	54	12.4%	27.8%
	White Male	510	303	162	31.8%	53.5%
	Black Male	240	100	30	12.5%	30.0%
	Hispanic Male	221	89	24	10.9%	27.0%
	White Female	471	324	202	42.9%	62.3%
	Black Female	244	152	50	20.5%	32.9%
	Hispanic Female	216	105	30	13.9%	28.6%
	Total	1,902	1,073	498	26.18%	46.4%

Notes: Number that Attempt Degree includes those who reported their highest grade completed as 1 year of college or more.

Table A.8: College Completion Expectations by Parent Education

		Parent's Degree	N	Mean	Median	Std. Deviation	Percent that Complete College
1997		Both HS or Less	1843	64.28%	75%	34.35	12.0%
		At Least 1 has Some College	1040	78.77%	90%	27.93	30.5%
		Both Have Only Some College	177	81.44%	90%	24.78	33.9%
		At Least 1 has a College Degree	810	86.42%	98%	22.05	52.5%
		Both Have at Least College Degree	303	90.25%	100%	18.44	67.7%
2001		Both HS or Less	970	56.27%	58.5%	39.99	12.4%
		At Least 1 has Some College	622	74.26%	98%	35.92	31.7%
		Both Have Only Some College	129	79.77%	100%	31.38	31.8%
		At Least 1 has a College Degree	453	87.70%	100%	25.16	57.2%
		Both Have at Least College Degree	181	94.33%	100%	14.80	72.9%

Notes: Education level is for genetic parents as of 1997. Does not take into account if both of the parents are in the household or not when interviewed.

Table A.9: College Completion Statistics by Parent Education

		Number of Students	Number that Attempt Degree	Number that Complete Degree	Percent that Complete Degree	Percent that Complete Degree Conditional on 1 Year Completion
1997	Both HS or Less	1,843	714	221	12.0%	31.0%
	At Least 1 has Some College	1,040	668	317	30.5%	47.5%
	Both Have Only Some College	177	127	60	33.9%	47.2%
	At Least 1 has a College Degree	810	639	425	52.5%	66.5%
	Both Have at Least College Degree	303	267	54	67.7%	76.8%
2001	Both HS or Less	970	390	120	12.4%	30.8%
	At Least 1 has Some College	622	422	197	31.7%	46.7%
	Both Have Only Some College	129	87	41	31.8%	47.1%
	At Least 1 has a College Degree	453	386	259	57.2%	67.1%
	Both Have at Least College Degree	181	169	132	72.9%	78.1%

Notes: Number that Attempt Degree includes those who reported their highest grade completed as 1 year of college or more.

Table A.10: ASVAB Percentile and College Degree Attainment

ASVAB Percentile	Number of Students	Percent that Attempt Degree	Percent that Complete Degree
0 - 25%	2,161	29.4 %	5.8 %
26 - 50 %	2,801	49.2 %	16.6 %
51 - 75 %	2,346	64.2 %	30.3 %
76 - 90 %	889	84.6 %	55.3 %
91 - 100 %	576	92.4 %	70.7 %

Notes: Number that Attempt Degree includes those who reported their highest grade completed as 1 year of college or more.

Table A.11: College Completion Expectations by ASVAB Percentile

	ASVAB Percentile	Number of Students	Mean Expectation	Percent that Attempt Degree	Percent that Complete Degree
1997 Sub-Sample	0 - 25%	844	62.1 %	30.1 %	5.7%
	26 - 50 %	1,100	69.0 %	49.2 %	17.5%
	51 - 75 %	929	75.7 %	62.8 %	30.4%
	76 - 90 %	348	88.1 %	84.5 %	53.7%
	91 - 100 %	277	90.7 %	93.0 %	73.1%
2001 Sub-Sample	0 - 25%	475	50.4 %	29.1 %	6.3%
	26 - 50 %	582	63.1 %	51.0 %	16.2%
	51 - 75 %	493	73.6 %	68.2 %	35.5%
	76 - 90 %	205	87.7 %	89.3 %	58.0%
	91 - 100 %	145	88.9 %	91.7 %	71.0%

Notes: Number that Attempt Degree includes those who reported their highest grade completed as 1 year of college or more.

Table A.12: Model and Variable Descriptions

Dependent Variables:	i. Student completes at least 1 year of college ii. Student completes a 4 year college degree	
Variable of Interest:	The percent chance the student believes that they will obtain a 4 year college degree by the time they are 30	
Control Variables:	Demographics - Male - White - Black Family - Only child - Household members under 18 - Mother, Father, or both parents absent - Income quantile Location - Urban - MSA Central City - MSA Non Central City - 3 Census Tracks (W, NE, NC)	Parent Education - One parent has some college - Both parents have some college - One parent has a college degree - Both parents have a college degree Schooling: - Self-reported grades in 8th grade (levels) - Self-reported grades in HS if applicable (levels) - ASVAB Percentile - Enrollment in high school indicator Other: - % of peers that plan to go to college - Teacher quality and involvement indicators

Notes: For each categorical variable, that which is not listed is used as the base case. For grade controls, if a student is still in high school then the 8th grade controls are used, when they have graduated or left school the high school grade controls are used. In each case the student reported their grades in levels. Transcript data for high school GPA is available but low reporting severely reduces sample size. Estimates from specifications that have included it do not differ drastically. The teacher quality and involvement controls are reported in levels by the student. The student reports whether they believe their teachers overall were "good" and whether they were "interested in the students success." They responded with strongly agree, agree, disagree, or strongly disagree.

Table A.13: College Expectations, Attendance, and Completion - Probit Model

	Y = 1 if completed at least 1 year of college			Y = 1 if completed a 4 year college degree		
	(1)	(2)	(3)	(4)	(5)	(6)
College Expectations _97	0.00699*** (0.000312)		0.00373*** (0.000350)	0.00589*** (0.000300)		0.00225*** (0.000303)
Demographic Controls	No	Yes	Yes	No	Yes	Yes
Parent Education Controls	No	Yes	Yes	No	Yes	Yes
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade & ASVAB	No	Yes	Yes	No	Yes	Yes
Peer and Teacher Controls	No	Yes	Yes	No	Yes	Yes
Observations	3450	3403	3403	3450	3366	3366
Pseudo R^2	0.125	0.261	0.286	0.123	0.326	0.340

Notes: Demographic controls include male, white, and black. Parent education include indicators for if one or both of the respondents' parents have some college or a four-year degree. Location controls include rural, 1 of 4 census tracks, and MSA. Family controls include number of members in the household under 18, whether student is only child, and indicators for if one or both parents are missing. Peer controls is the reported percent of peers that the respondent expects to go to college.

-Marginal effects; Robust standard errors in parentheses;

-* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.14: Expectations and Attendance - Linear Probability Model

Y = 1 if completed at least 1 year of college						
	1997 Sample			2001 High School Sample		
College Expectations_97	0.00603*** (0.000226)		0.00294*** (0.000264)			
College Expectations_01				0.00696*** (0.000333)		0.00435*** (0.000427)
Male		-0.0837*** (0.0150)	-0.0694*** (0.0148)		-0.0788** (0.0298)	-0.0506 (0.0277)
White		-0.0769** (0.0235)	-0.0661** (0.0231)		-0.0228 (0.0416)	-0.00820 (0.0384)
Black		0.0891*** (0.0242)	0.0636** (0.0239)		0.124** (0.0445)	0.0798 (0.0423)
Lngrossfaminc_1997		0.0287** (0.00938)	0.0244** (0.00908)		0.0197 (0.0176)	0.0161 (0.0168)
One parent some college		0.0614*** (0.0186)	0.0472** (0.0183)		0.0857* (0.0410)	0.0449 (0.0378)
Both parents some college		0.101** (0.0352)	0.0872* (0.0351)		0.107 (0.0765)	0.0430 (0.0744)
One parent 4yr degree		0.141*** (0.0216)	0.113*** (0.0215)		0.158** (0.0515)	0.102* (0.0478)
Both parents 4yr degree		0.0643* (0.0279)	0.0581* (0.0274)		0.0119 (0.0750)	-0.0257 (0.0675)
ASVAB Percentile		0.00497*** (0.000357)	0.00458*** (0.000354)		0.00535*** (0.000797)	0.00417*** (0.000728)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	3450	3450	3450	792	792	792
Adjusted R^2	0.149	0.277	0.303	0.300	0.332	0.427

Notes: The 2001 High School Sample are those who reported still being enrolled in high school when surveyed in 2001.

Robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.15: Expectations and Attendance - Probit Model

Y = 1 if completed at least 1 year of college						
	1997 Sample			2001 High School Sample		
College Expectations_97	0.0166*** (0.000766)		0.00886*** (0.000856)			
College Expectations_01				0.0219*** (0.00159)		0.0166*** (0.00183)
Male		-0.277*** (0.0494)	-0.239*** (0.0502)		-0.306** (0.112)	-0.229 (0.117)
White		-0.285*** (0.0777)	-0.246** (0.0785)		-0.118 (0.156)	-0.0586 (0.161)
Black		0.286*** (0.0767)	0.221** (0.0783)		0.474** (0.166)	0.366* (0.177)
Income Quantile		0.104*** (0.0298)	0.0935** (0.0302)		0.0964 (0.0707)	0.0938 (0.0736)
One parent some college		0.166** (0.0593)	0.131* (0.0596)		0.304* (0.138)	0.164 (0.148)
Both parents some college		0.303* (0.125)	0.267* (0.127)		0.304 (0.250)	0.110 (0.257)
One parent 4yr degree		0.435*** (0.0755)	0.348*** (0.0758)		0.583** (0.200)	0.354 (0.202)
Both parents 4yr degree		0.356** (0.133)	0.342* (0.134)		-0.0149 (0.364)	-0.174 (0.349)
ASVAB Percentile		0.0166*** (0.00129)	0.0157*** (0.00131)		0.0186*** (0.00294)	0.0162*** (0.00293)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	3450	3450	3450	792	792	792
Pseudo R^2	0.114	0.242	0.264	0.250	0.330	0.414

Notes: The 2001 High School Sample are those who reported still being enrolled in high school when surveyed in 2001.

Marginal Effects; Robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.16: Expectations and Matriculation - OLS

	Y = number of years of college completed		
College Expectations_97	0.0225*** (0.000706)		0.00911*** (0.000759)
Male	-0.228*** (0.0475)		-0.183*** (0.0468)
White	-0.145* (0.0715)		-0.112 (0.0704)
Black	0.415*** (0.0719)		0.336*** (0.0711)
Lngrossfaminc_1997	0.106*** (0.0296)		0.0926** (0.0287)
One parent some college	0.157** (0.0605)		0.113 (0.0594)
Both parents some college	0.349** (0.118)		0.307** (0.116)
One parent 4yr degree	0.632*** (0.0748)		0.546*** (0.0748)
Both parents 4yr degree	0.341** (0.106)		0.322** (0.105)
ASVAB Percentile	0.0176*** (0.00116)		0.0164*** (0.00115)
Location Controls	No	Yes	Yes
Family Controls	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes
Observations	3450	3450	3450
Adjusted R^2	0.174	0.392	0.414

Notes:

Robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.17: Expectations and Matriculation - Tobit and Ordered Probit

	Y = number of years of college completed					
	Tobit	Tobit	Tobit	OProbit	OProbit	OProbit
College Expectations_97	0.0470*** (0.00180)		0.0212*** (0.00174)	0.0184*** (0.000753)		0.00963*** (0.000816)
Male		-0.469*** (0.0851)	-0.367*** (0.0837)		-0.229*** (0.0432)	-0.185*** (0.0437)
White		-0.311* (0.134)	-0.244 (0.131)		-0.153* (0.0657)	-0.120 (0.0661)
Black		0.694*** (0.140)	0.526*** (0.138)		0.356*** (0.0673)	0.290*** (0.0682)
Lngrossfaminc_1997		0.217** (0.0666)	0.191** (0.0642)		0.110*** (0.0330)	0.0997** (0.0328)
One parent some college		0.334** (0.103)	0.252* (0.0999)		0.147** (0.0515)	0.115* (0.0516)
Both parents some college		0.540** (0.177)	0.458** (0.172)		0.257** (0.0951)	0.228* (0.0951)
One parent 4yr degree		0.957*** (0.113)	0.788*** (0.112)		0.518*** (0.0623)	0.445*** (0.0628)
Both parents 4yr degree		0.359* (0.149)	0.307* (0.146)		0.330** (0.101)	0.309** (0.102)
ASVAB Percentile		0.0302*** (0.00200)	0.0275*** (0.00198)		0.0154*** (0.00105)	0.0145*** (0.00106)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	3450	3450	3450	3450	3450	3450
Pseudo R^2	0.066	0.137	0.152	0.079	0.182	0.197

Notes: The 2001 High School Sample are those who reported still being enrolled in high school when surveyed in 2001.

Marginal Effects; Robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.18: Expectations and Completion - Linear Probability Model

Y = 1 if completed a 4 year college degree						
	1997 Sample			2001 High School Sample		
College Expectations_97	0.00469*** (0.000194)		0.00147*** (0.000204)			
College Expectations_01				0.00431*** (0.000315)		0.00172*** (0.000350)
Male		-0.0447*** (0.0135)	-0.0375** (0.0134)		-0.0242 (0.0246)	-0.0140 (0.0244)
White		-0.00843 (0.0195)	-0.00299 (0.0194)		-0.0189 (0.0333)	-0.0175 (0.0327)
Black		0.101*** (0.0203)	0.0878*** (0.0203)		0.00237 (0.0351)	-0.0113 (0.0352)
Lngrossfaminc_1997		0.0266*** (0.00768)	0.0244** (0.00760)		0.00484 (0.0137)	0.00477 (0.0140)
One parent some college		0.0246 (0.0171)	0.0175 (0.0170)		0.0401 (0.0348)	0.0270 (0.0339)
Both parents some college		0.0996** (0.0363)	0.0930** (0.0360)		0.162* (0.0727)	0.138 (0.0738)
One parent 4yr degree		0.168*** (0.0219)	0.154*** (0.0221)		0.140** (0.0532)	0.119* (0.0529)
Both parents 4yr degree		0.117*** (0.0326)	0.114*** (0.0325)		0.217** (0.0824)	0.204* (0.0802)
ASVAB Percentile		0.00334*** (0.000337)	0.00314*** (0.000335)		0.00251*** (0.000730)	0.00216** (0.000706)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	3450	3450	3450	792	792	792
Adjusted R^2	0.107	0.307	0.315	0.164	0.350	0.366

Notes: The 2001 High School Sample are those who reported still being enrolled in high school when surveyed in 2001.

Robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.19: Expectations and Completion - Probit Model

Y = 1 if completed a 4 year college degree						
	1997 Sample			2001 High School Sample		
College Expectations_97	0.00578*** (0.000313)		0.00230*** (0.000332)			
College Expectations_01				0.00492*** (0.000403)		0.00281*** (0.000439)
Male (d)		-0.0570*** (0.0167)	-0.0443** (0.0166)		-0.0316 (0.0299)	-0.0136 (0.0281)
White (d)		-0.00743 (0.0255)	-0.000583 (0.0253)		-0.0340 (0.0410)	-0.0195 (0.0386)
Black (d)		0.134*** (0.0311)	0.114*** (0.0308)		0.00719 (0.0491)	-0.00433 (0.0446)
Lngrossfaminc_1997		0.0448** (0.0145)	0.0416** (0.0141)		0.00771 (0.0195)	0.000786 (0.0180)
One parent some college (d)		0.0302 (0.0203)	0.0214 (0.0198)		0.0558 (0.0399)	0.0259 (0.0351)
Both parents some college (d)		0.109* (0.0425)	0.0995* (0.0419)		0.162 (0.0844)	0.113 (0.0772)
One parent 4yr degree (d)		0.180*** (0.0252)	0.160*** (0.0249)		0.127* (0.0585)	0.0827 (0.0509)
Both parents 4yr degree (d)		0.113** (0.0399)	0.105** (0.0392)		0.236* (0.117)	0.191 (0.106)
ASVAB Percentile		0.00367*** (0.000394)	0.00334*** (0.000390)		0.00252*** (0.000720)	0.00180** (0.000646)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	3450	3450	3450	792	758	758
Pseudo R^2	0.105	0.277	0.289	0.193	0.343	0.383

Notes: The 2001 High School Sample are those who reported still being enrolled in high school when surveyed in 2001.

Marginal effects; Robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.20: College Expectations and Completion with 2001 Sample - Probit Model

	Y = 1 if completed a 4 year college degree					
	(1)	(2)	(3)	(4)	(5)	(6)
College Expectations_01	0.00719*** (0.000287)		0.00506*** (0.000325)	0.00719*** (0.000287)		0.00386*** (0.000361)
Enrolled in 2yr College (d)					0.258*** (0.0510)	0.149*** (0.0440)
Enrolled in 4yr College (d)					0.607*** (0.0369)	0.421*** (0.0471)
Demographic Controls	No	Yes	Yes	No	Yes	Yes
Parent Education Controls	No	Yes	Yes	No	Yes	Yes
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in High School	No	Yes	Yes	No	Yes	Yes
Peer and Teacher Controls	No	Yes	Yes	No	Yes	Yes
Observations	1902	1815	1815	1902	1815	1815
Pseudo R^2	0.254	0.334	0.433	0.254	0.448	0.492

Notes: Demographic controls include male, white, and black. Parent education include indicators for if one or both of the respondents' parents have some college or a four-year degree. Location controls include rural, 1 of 4 census tracks, and MSA. Family controls include number of members in the household under 18, whether student is only child, and indicators for if one or both parents are missing. Peer controls is the reported percent of peers that the respondent expects to go to college.

-Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

-* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.21: 2001 Expectations and Completion w/ Enrollment Controls

	Y = 1 if completed a 4 year college degree					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_01	0.00579*** (0.000197)		0.00214*** (0.000213)	0.00785*** (0.000381)		0.00428*** (0.000434)
Male (d)		-0.0360* (0.0161)	-0.0280 (0.0158)		-0.0610* (0.0255)	-0.0422 (0.0243)
White (d)		0.00569 (0.0223)	0.0196 (0.0222)		0.00336 (0.0358)	0.0268 (0.0342)
Black (d)		0.0159 (0.0247)	0.00680 (0.0243)		0.0267 (0.0423)	0.00534 (0.0393)
Lngrossfaminc_1997		0.00710 (0.00962)	0.00394 (0.00950)		0.00879 (0.0175)	-0.0000815 (0.0168)
One parent some college (d)		0.0192 (0.0197)	0.0108 (0.0193)		0.0207 (0.0291)	0.00799 (0.0276)
Both parents some college (d)		0.100* (0.0413)	0.0797 (0.0409)		0.152* (0.0601)	0.117* (0.0580)
One parent 4yr degree (d)		0.126*** (0.0278)	0.104*** (0.0275)		0.170*** (0.0391)	0.130*** (0.0372)
Both parents 4yr degree (d)		0.117** (0.0402)	0.110** (0.0391)		0.170* (0.0681)	0.134* (0.0625)
ASVAB Percentile		0.00140*** (0.000421)	0.00101* (0.000411)		0.00174** (0.000603)	0.00105 (0.000573)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Enrollment Controls	No	Yes	Yes	No	Yes	Yes
Observations	1902	1902	1902	1902	1850	1850
Adjusted R^2	0.230	0.465	0.486			
Pseudo R^2				0.242	0.425	0.465

Notes: The 2001 High School Sample are those who reported still being enrolled in high school when surveyed in 2001.

Robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.22: 2001 Expectations and Completion w/ High School Grade Controls

	Y = 1 if completed a 4 year college degree					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_01	0.00624*** (0.000224)		0.00211*** (0.000255)	0.00870*** (0.000412)		0.00489*** (0.000508)
Male (d)		-0.0340 (0.0185)	-0.0449* (0.0180)		-0.0667* (0.0324)	-0.0737** (0.0281)
White (d)		0.0149 (0.0261)	0.0359 (0.0262)		0.0170 (0.0450)	0.0612 (0.0394)
Black (d)		0.0268 (0.0290)	0.0203 (0.0289)		0.0504 (0.0544)	0.0241 (0.0468)
Lngrossfaminc_1997		0.000725 (0.0105)	-0.00212 (0.0101)		0.00840 (0.0242)	0.000952 (0.0192)
One parent some college (d)		-0.00842 (0.0224)	-0.00177 (0.0221)		-0.0198 (0.0361)	-0.0000922 (0.0317)
Both parents some college (d)		0.0917* (0.0466)	0.0554 (0.0462)		0.170* (0.0810)	0.0884 (0.0689)
One parent 4yr degree (d)		0.0987** (0.0315)	0.0800* (0.0314)		0.157** (0.0482)	0.106* (0.0418)
Both parents 4yr degree (d)		0.0738 (0.0457)	0.0849 (0.0453)		0.112 (0.0803)	0.0972 (0.0708)
ASVAB Percentile		0.00101* (0.000489)	0.00140** (0.000469)		0.00169* (0.000783)	0.00201** (0.000693)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in High School	No	Yes	No	No	Yes	No
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Enrollment Controls	No	Yes	Yes	No	Yes	Yes
Observations	1311	1311	1311	1311	1271	1311
Adjusted R^2	0.274	0.532	0.540			
Pseudo R^2				0.294	0.482	0.517

Notes:

Marginal effects; Robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.23: College Completion of Overlapping Sample in 1997 and 2001 - Probit Model

Y = 1 if completed a 4 year college degree							
	1997 Sample			1997 Sample 5 Years Later			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
College Expectations_97	0.00686*** (0.000704)		0.00290*** (0.000754)				0.000296 (0.000599)
College Expectations_01				0.00700*** (0.000559)		0.00380*** (0.000569)	0.00378*** (0.000584)
Enrolled in 2yr College (d)					0.289*** (0.0715)	0.157* (0.0621)	0.157* (0.0624)
Enrolled in 4yr College (d)					0.645*** (0.0494)	0.391*** (0.0683)	0.389*** (0.0690)
Demographic Controls	No	Yes	Yes	No	Yes	Yes	Yes
Parent Education Controls	No	Yes	Yes	No	Yes	Yes	Yes
Location Controls	No	Yes	Yes	No	Yes	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes	Yes
Grades in High School	No	No	No	No	Yes	Yes	Yes
Peer and Teacher Controls	No	Yes	Yes	No	Yes	Yes	Yes
Observations	700	663	663	700	651	651	651
Pseudo R^2	0.135	0.355	0.370	0.299	0.551	0.595	0.596

Notes: Demographic controls include male, white, and black. Parent education include indicators for if one or both of the respondents' parents have some college or a four-year degree. Location controls include rural, 1 of 4 census tracts, and MSA. Family controls include number of members in the household under 18, whether student is only child, and indicators for if one or both parents are missing. Peer controls is the reported percent of peers that the respondent expects to go to college.

-Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1
 - * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.24: 1997 Expectations and Attendance - Race Interactions

	Y = 1 if completed at least 1 year of college					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_97	0.00555*** (0.000339)		0.00226*** (0.000561)	0.00611*** (0.000404)		0.00237** (0.000734)
Male Expectations_97	-0.000953*** (0.000199)		0.000122 (0.000434)	-0.00107*** (0.000226)		0.000376 (0.000632)
White Expectations_97	0.00150*** (0.000279)		0.00104 (0.000571)	0.00168*** (0.000303)		0.00160* (0.000803)
Black Expectations_97	-0.0000427 (0.000316)		0.000384 (0.000631)	-0.0000659 (0.000331)		0.000592 (0.000859)
Male (d)		-0.0837*** (0.0150)	-0.0773* (0.0349)		-0.107*** (0.0191)	-0.118* (0.0493)
White (d)		-0.0769** (0.0235)	-0.138** (0.0442)		-0.106*** (0.0299)	-0.202*** (0.0610)
Black (d)		0.0891*** (0.0242)	0.0393 (0.0484)		0.112*** (0.0289)	0.0492 (0.0667)
Ln(GrossFamIncome)		0.0287** (0.00938)	0.0247** (0.00909)		0.0353** (0.0128)	0.0317* (0.0128)
One parent some college (d)		0.0614*** (0.0186)	0.0473** (0.0183)		0.0689** (0.0227)	0.0552* (0.0231)
Both parents some college (d)		0.101** (0.0352)	0.0865* (0.0351)		0.118** (0.0442)	0.104* (0.0458)
One parent 4yr degree (d)		0.141*** (0.0216)	0.111*** (0.0215)		0.172*** (0.0267)	0.137*** (0.0277)
Both parents 4yr degree (d)		0.0643* (0.0279)	0.0561* (0.0274)		0.137** (0.0467)	0.130** (0.0477)
ASVAB Percentile		0.00497*** (0.000357)	0.00457*** (0.000354)		0.00651*** (0.000502)	0.00618*** (0.000512)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	3450	3450	3450	3450	3450	3450
Adjusted R^2	0.169	0.277	0.303			
Pseudo R^2				0.131	0.241	0.265

Notes:

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1
 * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.25: 2001 Expectations and Attendance - Race Interactions

	Y = 1 if completed at least 1 year of college					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_01	0.00624*** (0.000594)		0.00356*** (0.000752)	0.00770*** (0.000774)		0.00536*** (0.00134)
Male Expectations_01	-0.000494 (0.000436)		-0.000383 (0.000607)	-0.000520 (0.000505)		-0.000216 (0.00129)
White Expectations_01	0.00193*** (0.000567)		0.000592 (0.000800)	0.00214*** (0.000635)		0.000845 (0.00164)
Black Expectations_01	-0.000241 (0.000668)		-0.0000512 (0.000867)	-0.000259 (0.000686)		-0.000408 (0.00170)
Male (d)		-0.0724* (0.0286)	-0.0278 (0.0401)		-0.128** (0.0445)	-0.0797 (0.0995)
White (d)		0.00491 (0.0401)	-0.0252 (0.0531)		0.00562 (0.0638)	-0.0398 (0.133)
Black (d)		0.0996* (0.0425)	0.0763 (0.0524)		0.167* (0.0669)	0.169 (0.139)
Ln(GrossFamIncome)		0.0109 (0.0165)	0.0103 (0.0162)		0.0201 (0.0265)	0.0170 (0.0276)
One parent some college (d)		0.0658 (0.0393)	0.0384 (0.0372)		0.105 (0.0553)	0.0605 (0.0572)
Both parents some college (d)		0.0957 (0.0708)	0.0440 (0.0705)		0.106 (0.0962)	0.0476 (0.0964)
One parent 4yr degree (d)		0.154** (0.0473)	0.104* (0.0464)		0.250*** (0.0757)	0.161* (0.0792)
Both parents 4yr degree (d)		-0.00211 (0.0710)	-0.0291 (0.0662)		-0.0442 (0.131)	-0.0826 (0.120)
ASVAB Percentile		0.00460*** (0.000775)	0.00385*** (0.000724)		0.00651*** (0.00117)	0.00580*** (0.00116)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	792	792	792	792	792	792
Adjusted R^2	0.321	0.400	0.447			
Pseudo R^2				0.269	0.389	0.444

Notes:

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.26: 1997 Expectations and Completion - Race Interactions

	Y = 1 if completed a 4 year college degree					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_97	0.00351*** (0.000293)		0.000559 (0.000440)	0.00460*** (0.000392)		0.00116 (0.000777)
Male Expectations_97	-0.000739*** (0.000202)		-0.000157 (0.000358)	-0.000624*** (0.000185)		0.000377 (0.000641)
White Expectations_97	0.00233*** (0.000259)		0.00156*** (0.000460)	0.00209*** (0.000259)		0.00147 (0.000855)
Black Expectations_97	0.000455 (0.000284)		0.000842 (0.000501)	0.000420 (0.000288)		0.000726 (0.000929)
Male (d)		-0.0447*** (0.0135)	-0.0249 (0.0251)		-0.0570*** (0.0167)	-0.0734 (0.0543)
White (d)		-0.00843 (0.0195)	-0.111*** (0.0319)		-0.00743 (0.0255)	-0.115 (0.0718)
Black (d)		0.101*** (0.0203)	0.0321 (0.0344)		0.134*** (0.0311)	0.0554 (0.0827)
Ln(GrossFamIncome)		0.0266*** (0.00768)	0.0248** (0.00756)		0.0448** (0.0145)	0.0418** (0.0140)
One parent some college (d)		0.0246 (0.0171)	0.0178 (0.0170)		0.0302 (0.0203)	0.0220 (0.0199)
Both parents some college (d)		0.0996** (0.0363)	0.0922* (0.0358)		0.109* (0.0425)	0.0985* (0.0417)
One parent 4yr degree (d)		0.168*** (0.0219)	0.152*** (0.0221)		0.180*** (0.0252)	0.158*** (0.0249)
Both parents 4yr degree (d)		0.117*** (0.0326)	0.110*** (0.0326)		0.113** (0.0399)	0.103** (0.0394)
ASVAB Percentile		0.00334*** (0.000337)	0.00312*** (0.000336)		0.00367*** (0.000394)	0.00334*** (0.000392)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	3450	3450	3450	3450	3450	3450
Adjusted R^2	0.143	0.307	0.316			
Pseudo R^2				0.132	0.277	0.290

Notes:

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.27: 2001 Expectations and Completion - Race Interactions

	Y = 1 if completed a 4 year college degree					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_01	0.00330*** (0.000547)		0.00112 (0.000652)	0.00413*** (0.000528)		0.000965 (0.000838)
Male Expectations_01	-0.0000685 (0.000445)		0.000471 (0.000558)	-0.0000406 (0.000309)		0.00150 (0.000903)
White Expectations_01	0.00245*** (0.000571)		0.00154* (0.000757)	0.00154*** (0.000410)		0.00219 (0.00118)
Black Expectations_01	-0.000919 (0.000584)		-0.000783 (0.000753)	-0.000708 (0.000463)		-0.000480 (0.00111)
Male (d)		-0.0242 (0.0246)	-0.0412 (0.0281)		-0.0293 (0.0288)	-0.134 (0.0796)
White (d)		-0.0189 (0.0333)	-0.104* (0.0405)		-0.0264 (0.0409)	-0.182 (0.0959)
Black (d)		0.00237 (0.0351)	0.0399 (0.0381)		-0.000983 (0.0458)	0.0367 (0.100)
Ln(GrossFamIncome)		0.00484 (0.0137)	0.00463 (0.0139)		0.00456 (0.0180)	0.00208 (0.0161)
One parent some college (d)		0.0401 (0.0348)	0.0274 (0.0338)		0.0441 (0.0373)	0.0195 (0.0330)
Both parents some college (d)		0.162* (0.0727)	0.126 (0.0734)		0.147 (0.0799)	0.100 (0.0730)
One parent 4yr degree (d)		0.140** (0.0532)	0.0967 (0.0526)		0.118* (0.0558)	0.0624 (0.0469)
Both parents 4yr degree (d)		0.217** (0.0824)	0.200* (0.0785)		0.216 (0.112)	0.173 (0.101)
ASVAB Percentile		0.00251*** (0.000730)	0.00210** (0.000703)		0.00214** (0.000687)	0.00163** (0.000616)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	792	792	792	792	758	758
Adjusted R^2	0.229	0.350	0.372			
Pseudo R^2				0.242	0.358	0.399

Notes:

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.28: 1997 Expectations and Attendance - Gender with Race Interactions

	Y = 1 if completed at least 1 year of college					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_97	0.00565*** (0.000403)		0.00233*** (0.000595)	0.00618*** (0.000470)		0.00239** (0.000771)
White Male Expectations_97	0.000752 (0.000396)		0.00134 (0.000738)	0.000843 (0.000435)		0.00227* (0.00104)
White Female Expectations_97	0.00117** (0.000384)		0.000857 (0.000606)	0.00136** (0.000430)		0.00139 (0.000848)
Black Male Expectations_97	-0.00159*** (0.000457)		0.0000497 (0.000815)	-0.00165*** (0.000478)		0.000541 (0.00110)
Black Female Expectations_97	0.000284 (0.000430)		0.000593 (0.000670)	0.000298 (0.000470)		0.000896 (0.000913)
Hispanic Male Expectations_97	-0.00114* (0.000495)		-0.0000131 (0.000635)	-0.00117* (0.000517)		0.000330 (0.000826)
Male (d)		-0.0837*** (0.0150)	-0.0762* (0.0350)		-0.107*** (0.0191)	-0.119* (0.0494)
White (d)		-0.0769** (0.0235)	-0.142** (0.0443)		-0.106*** (0.0299)	-0.206*** (0.0611)
Black (d)		0.0891*** (0.0242)	0.0409 (0.0485)		0.112*** (0.0289)	0.0503 (0.0668)
Ln(GrossFamIncome)		0.0287** (0.00938)	0.0250** (0.00913)		0.0353** (0.0128)	0.0321* (0.0128)
One parent some college (d)		0.0614*** (0.0186)	0.0476** (0.0183)		0.0689** (0.0227)	0.0557* (0.0231)
Both parents some college (d)		0.101** (0.0352)	0.0875* (0.0352)		0.118** (0.0442)	0.104* (0.0458)
One parent 4yr degree (d)		0.141*** (0.0216)	0.111*** (0.0215)		0.172*** (0.0267)	0.137*** (0.0277)
Both parents 4yr degree (d)		0.0643* (0.0279)	0.0558* (0.0274)		0.137** (0.0467)	0.129** (0.0477)
ASVAB Percentile		0.00497*** (0.000357)	0.00453*** (0.000354)		0.00651*** (0.000502)	0.00615*** (0.000512)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	3450	3450	3450	3450	3450	3450
Adjusted R^2	0.171	0.277	0.304			
Pseudo R^2				0.132	0.241	0.266

Notes:

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.29: 2001 Expectations and Attendance - Gender with Race Interactions

	Y = 1 if completed at least 1 year of college					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_01	0.00561*** (0.000718)		0.00298*** (0.000860)	0.00707*** (0.000859)		0.00468*** (0.00142)
White Male Expectations_01	0.00232** (0.000778)		0.000950 (0.00103)	0.00259** (0.000856)		0.00167 (0.00200)
White Female Expectations_01	0.00235** (0.000778)		0.00109 (0.000956)	0.00253** (0.000861)		0.00126 (0.00177)
Black Male Expectations_01	-0.00111 (0.000935)		-0.000612 (0.00114)	-0.00114 (0.000943)		-0.000782 (0.00204)
Black Female Expectations_01	0.00140 (0.000904)		0.00126 (0.00104)	0.00148 (0.000959)		0.00140 (0.00189)
Hispanic Male Expectations_01	0.000965 (0.000972)		0.000917 (0.000976)	0.00102 (0.00102)		0.00144 (0.00161)
Male (d)		-0.0724* (0.0286)	-0.0284 (0.0402)		-0.128** (0.0445)	-0.0853 (0.100)
White (d)		0.00491 (0.0401)	-0.0241 (0.0530)		0.00562 (0.0638)	-0.0386 (0.134)
Black (d)		0.0996* (0.0425)	0.0761 (0.0523)		0.167* (0.0669)	0.160 (0.140)
Ln(GrossFamIncome)		0.0109 (0.0165)	0.00751 (0.0161)		0.0201 (0.0265)	0.0123 (0.0277)
One parent some college (d)		0.0658 (0.0393)	0.0390 (0.0368)		0.105 (0.0553)	0.0655 (0.0577)
Both parents some college (d)		0.0957 (0.0708)	0.0460 (0.0710)		0.106 (0.0962)	0.0460 (0.0975)
One parent 4yr degree (d)		0.154** (0.0473)	0.107* (0.0463)		0.250*** (0.0757)	0.168* (0.0800)
Both parents 4yr degree (d)		-0.00211 (0.0710)	-0.0347 (0.0663)		-0.0442 (0.131)	-0.0974 (0.119)
ASVAB Percentile		0.00460*** (0.000775)	0.00373*** (0.000730)		0.00651*** (0.00117)	0.00566*** (0.00116)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	792	792	792	792	792	792
Adjusted R^2	0.328	0.400	0.451			
Pseudo R^2				0.276	0.389	0.449

Notes:

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.30: 1997 Expectations and Completion - Gender with Race Interactions

	Y = 1 if completed a 4 year college degree					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_97	0.00378*** (0.000367)		0.000806 (0.000485)	0.00494*** (0.000433)		0.00149 (0.000791)
White Male Expectations_97	0.00150*** (0.000388)		0.00128* (0.000614)	0.00134*** (0.000356)		0.00167 (0.00107)
White Female Expectations_97	0.00190*** (0.000381)		0.00126* (0.000516)	0.00159*** (0.000349)		0.00109 (0.000879)
Black Male Expectations_97	-0.000714 (0.000417)		0.000376 (0.000667)	-0.000733 (0.000410)		0.000714 (0.00114)
Black Female Expectations_97	0.000315 (0.000418)		0.000645 (0.000563)	0.000239 (0.000386)		0.000436 (0.000954)
Hispanic Male Expectations_97	-0.00129** (0.000428)		-0.000710 (0.000513)	-0.00135** (0.000462)		-0.000467 (0.000796)
Male (d)		-0.0447*** (0.0135)	-0.0222 (0.0252)		-0.0570*** (0.0167)	-0.0689 (0.0545)
White (d)		-0.00843 (0.0195)	-0.116*** (0.0319)		-0.00743 (0.0255)	-0.123 (0.0718)
Black (d)		0.101*** (0.0203)	0.0294 (0.0344)		0.134*** (0.0311)	0.0508 (0.0822)
Ln(GrossFamIncome)		0.0266*** (0.00768)	0.0250*** (0.00757)		0.0448** (0.0145)	0.0421** (0.0141)
One parent some college (d)		0.0246 (0.0171)	0.0181 (0.0170)		0.0302 (0.0203)	0.0224 (0.0199)
Both parents some college (d)		0.0996** (0.0363)	0.0929** (0.0358)		0.109* (0.0425)	0.0999* (0.0415)
One parent 4yr degree (d)		0.168*** (0.0219)	0.152*** (0.0221)		0.180*** (0.0252)	0.159*** (0.0249)
Both parents 4yr degree (d)		0.117*** (0.0326)	0.110*** (0.0326)		0.113** (0.0399)	0.101** (0.0393)
ASVAB Percentile		0.00334*** (0.000337)	0.00311*** (0.000336)		0.00367*** (0.000394)	0.00334*** (0.000392)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	3450	3450	3450	3450	3450	3450
Adjusted R^2	0.143	0.307	0.316			
Pseudo R^2				0.133	0.277	0.291

Notes:

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.31: 2001 Expectations and Completion - Gender with Race Interactions

	Y = 1 if completed a 4 year college degree					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_01	0.00331*** (0.000653)		0.00122 (0.000720)	0.00413*** (0.000583)		0.00107 (0.000869)
White Male Expectations_01	0.00271*** (0.000779)		0.00208* (0.000938)	0.00173** (0.000541)		0.00373** (0.00123)
White Female Expectations_01	0.00215** (0.000782)		0.00134 (0.000855)	0.00131* (0.000537)		0.00202 (0.00121)
Black Male Expectations_01	-0.00156* (0.000760)		-0.000597 (0.000951)	-0.00136* (0.000670)		0.000718 (0.00145)
Black Female Expectations_01	-0.000346 (0.000833)		-0.000684 (0.000906)	-0.000216 (0.000615)		-0.000403 (0.00115)
Hispanic Male Expectations_01	-0.0000579 (0.000919)		0.000266 (0.000934)	-0.0000457 (0.000662)		0.00129 (0.00109)
Male (d)		-0.0242 (0.0246)	-0.0396 (0.0282)		-0.0293 (0.0288)	-0.132 (0.0802)
White (d)		-0.0189 (0.0333)	-0.104* (0.0406)		-0.0264 (0.0409)	-0.182 (0.0964)
Black (d)		0.00237 (0.0351)	0.0390 (0.0382)		-0.000983 (0.0458)	0.0354 (0.100)
Ln(GrossFamIncome)		0.00484 (0.0137)	0.00389 (0.0138)		0.00456 (0.0180)	0.00137 (0.0161)
One parent some college (d)		0.0401 (0.0348)	0.0268 (0.0337)		0.0441 (0.0373)	0.0184 (0.0328)
Both parents some college (d)		0.162* (0.0727)	0.126 (0.0738)		0.147 (0.0799)	0.101 (0.0736)
One parent 4yr degree (d)		0.140** (0.0532)	0.0969 (0.0526)		0.118* (0.0558)	0.0628 (0.0470)
Both parents 4yr degree (d)		0.217** (0.0824)	0.195* (0.0787)		0.216 (0.112)	0.166 (0.0997)
ASVAB Percentile		0.00251*** (0.000730)	0.00206** (0.000711)		0.00214** (0.000687)	0.00160** (0.000619)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	792	792	792	792	758	758
Adjusted R^2	0.231	0.350	0.371			
Pseudo R^2				0.246	0.358	0.400

Notes:

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.32: Quadratic and Cubic Expectations

	<u>Y = 1 if completed at least 1 year of college</u>		<u>Y = 1 if completed a 4 year college degree</u>			
	<u>1997 Sample</u>		<u>1997 Sample</u>		<u>2001 Sample</u>	
	(1)	(2)	(3)	(4)	(5)	(6)
College Expectations_97	0.00523*** (0.00138)	0.000898 (0.00341)	0.00212 (0.00132)	-0.000516 (0.00324)		
Expectations ²	-0.0000129 (0.0000115)	0.0000926 (0.0000781)	0.00000107 (0.0000104)	0.0000608 (0.0000682)		
Expectations ³		-0.000000655 (0.000000482)		-0.000000354 (0.000000400)		
College Expectations_01					0.00272 (0.00201)	-0.00261 (0.00514)
Expectations ²					0.0000126 (0.0000171)	0.000135 (0.000108)
Expectations ³						-0.000000726 (0.000000624)
Demographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Parent Education Controls	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes
Family Controls	Yes	Yes	Yes	Yes	Yes	Yes
Grades in 8th Grade	Yes	Yes	Yes	Yes	No	No
Grades in High School	No	No	No	No	Yes	Yes
Peer and Teacher Controls	Yes	Yes	Yes	Yes	Yes	Yes
College Enrollment Controls	No	No	No	No	Yes	Yes
Observations	3403	3403	3366	3366	1815	1815
Pseudo R^2	0.286	0.286	0.340	0.340	0.484	0.485

Notes: Demographic controls include male, white, and black. Parent education include indicators for if one or both of the respondents' parents have some college or a four-year degree. Location controls include rural, 1 of 4 census tracks, and MSA. Family controls include number of members in the household under 18, whether student is only child, and indicators for if one or both parents are missing. Peer controls is the reported percent of peers that the respondent expects to go to college.

-Marginal effects; Robust standard errors in parentheses

—* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.33: 1997 Expectations and Attendance - Quadratic

	Y = 1 if completed at least 1 year of college					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_97	0.00580** (0.00192)		0.00350*** (0.000893)	0.00529*** (0.00125)		0.00491*** (0.00132)
Expectations ²	0.00000909 (0.0000177)		-0.00000503 (0.00000798)	0.0000110 (0.0000103)		-0.0000123 (0.0000110)
Male (d)		-0.0837*** (0.0150)	-0.0700*** (0.0149)		-0.107*** (0.0191)	-0.0936*** (0.0196)
White (d)		-0.0769** (0.0235)	-0.0657** (0.0231)		-0.106*** (0.0299)	-0.0915** (0.0304)
Black (d)		0.0891*** (0.0242)	0.0645** (0.0240)		0.112*** (0.0289)	0.0891** (0.0300)
Lngrossfaminc_1997		0.0287** (0.00938)	0.0244** (0.00908)		0.0353** (0.0128)	0.0318* (0.0129)
One parent some college (d)		0.0614*** (0.0186)	0.0474** (0.0183)		0.0689** (0.0227)	0.0557* (0.0230)
Both parents some college (d)		0.101** (0.0352)	0.0869* (0.0351)		0.118** (0.0442)	0.104* (0.0457)
One parent 4yr degree (d)		0.141*** (0.0216)	0.114*** (0.0215)		0.172*** (0.0267)	0.140*** (0.0276)
Both parents 4yr degree (d)		0.0643* (0.0279)	0.0587* (0.0275)		0.137** (0.0467)	0.133** (0.0472)
ASVAB Percentile		0.00497*** (0.000357)	0.00458*** (0.000354)		0.00651*** (0.000502)	0.00618*** (0.000511)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	3450	3450	3450	3450	3450	3450
Adjusted R^2	0.156	0.277	0.303			
Pseudo R^2				0.114	0.241	0.264

Notes:

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.34: 2001 Expectations and Attendance - Quadratic

	Y = 1 if completed at least 1 year of college					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_01	-0.000804 (0.00148)		-0.00108 (0.00138)	0.00149 (0.00222)		0.00112 (0.00261)
Expectations ²	0.0000755*** (0.0000146)		0.0000465** (0.0000142)	0.0000648** (0.0000203)		0.0000405 (0.0000237)
Male (d)		-0.0724* (0.0286)	-0.0487 (0.0269)		-0.128** (0.0445)	-0.0957* (0.0458)
White (d)		0.00491 (0.0401)	0.0160 (0.0380)		0.00562 (0.0638)	0.0200 (0.0650)
Black (d)		0.0996* (0.0425)	0.0728 (0.0411)		0.167* (0.0669)	0.136 (0.0702)
Lngrossfaminc_1997		0.0109 (0.0165)	0.00680 (0.0164)		0.0201 (0.0265)	0.0135 (0.0282)
One parent some college (d)		0.0658 (0.0393)	0.0290 (0.0371)		0.105 (0.0553)	0.0566 (0.0574)
Both parents some college (d)		0.0957 (0.0708)	0.0370 (0.0704)		0.106 (0.0962)	0.0439 (0.0971)
One parent 4yr degree (d)		0.154** (0.0473)	0.0977* (0.0448)		0.250*** (0.0757)	0.165* (0.0782)
Both parents 4yr degree (d)		-0.00211 (0.0710)	-0.0363 (0.0646)		-0.0442 (0.131)	-0.0836 (0.119)
ASVAB Percentile		0.00460*** (0.000775)	0.00383*** (0.000720)		0.00651*** (0.00117)	0.00585*** (0.00116)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	792	792	792	792	792	792
Adjusted R^2	0.322	0.400	0.456			
Pseudo R^2				0.259	0.389	0.446

Notes: 2001 Sub-sample limited to those with either a HS Degree or older than 18 in 2001.

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.35: 1997 Expectations and Completion - Quadratic

	Y = 1 if completed a 4 year college degree					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_97	-0.00302 (0.00163)		-0.00110 (0.000704)	0.00140 (0.00135)		0.000662 (0.00135)
Expectations ²	0.0000776*** (0.0000160)		0.0000230*** (0.00000653)	0.0000358*** (0.0000107)		0.0000137 (0.0000109)
Male (d)		-0.0447*** (0.0135)	-0.0345** (0.0134)		-0.0570*** (0.0167)	-0.0428* (0.0167)
White (d)		-0.00843 (0.0195)	-0.00452 (0.0194)		-0.00743 (0.0255)	-0.000957 (0.0253)
Black (d)		0.101*** (0.0203)	0.0839*** (0.0203)		0.134*** (0.0311)	0.112*** (0.0308)
Lngrossfaminc_1997		0.0266*** (0.00768)	0.0244** (0.00759)		0.0448** (0.0145)	0.0417** (0.0141)
One parent some college (d)		0.0246 (0.0171)	0.0167 (0.0169)		0.0302 (0.0203)	0.0204 (0.0198)
Both parents some college (d)		0.0996** (0.0363)	0.0944** (0.0359)		0.109* (0.0425)	0.100* (0.0420)
One parent 4yr degree (d)		0.168*** (0.0219)	0.152*** (0.0220)		0.180*** (0.0252)	0.160*** (0.0249)
Both parents 4yr degree (d)		0.117*** (0.0326)	0.111*** (0.0324)		0.113** (0.0399)	0.103** (0.0392)
ASVAB Percentile		0.00334*** (0.000337)	0.00314*** (0.000335)		0.00367*** (0.000394)	0.00335*** (0.000391)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	3450	3450	3450	3450	3450	3450
Adjusted R^2	0.121	0.307	0.317			
Pseudo R^2				0.108	0.277	0.289

Notes:

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.36: 2001 Expectations and Completion - Quadratic

	Y = 1 if completed a 4 year college degree					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_01	-0.00285** (0.00109)		-0.00159 (0.00109)	0.000118 (0.00196)		-0.0000775 (0.00195)
Expectations ²	0.0000696*** (0.0000115)		0.0000330** (0.0000117)	0.0000433* (0.0000173)		0.0000232 (0.0000174)
Male (d)		-0.0242 (0.0246)	-0.0122 (0.0243)		-0.0293 (0.0288)	-0.0120 (0.0278)
White (d)		-0.0189 (0.0333)	-0.0118 (0.0323)		-0.0264 (0.0409)	-0.0119 (0.0383)
Black (d)		0.00237 (0.0351)	-0.0102 (0.0350)		-0.000983 (0.0458)	-0.00457 (0.0431)
Lngrossfaminc_1997		0.00484 (0.0137)	0.00192 (0.0140)		0.00456 (0.0180)	-0.00177 (0.0177)
One parent some college (d)		0.0401 (0.0348)	0.0201 (0.0338)		0.0441 (0.0373)	0.0190 (0.0338)
Both parents some college (d)		0.162* (0.0727)	0.131 (0.0737)		0.147 (0.0799)	0.108 (0.0743)
One parent 4yr degree (d)		0.140** (0.0532)	0.111* (0.0530)		0.118* (0.0558)	0.0796 (0.0501)
Both parents 4yr degree (d)		0.217** (0.0824)	0.198* (0.0792)		0.216 (0.112)	0.183 (0.103)
ASVAB Percentile		0.00251*** (0.000730)	0.00213** (0.000700)		0.00214** (0.000687)	0.00162* (0.000634)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	792	792	792	792	758	758
Adjusted R^2	0.189	0.350	0.371			
Pseudo R^2				0.201	0.358	0.391

Notes: 2001 Sub-sample limited to those with either a HS Degree or older than 18 in 2001.

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.37: 1997 Expectations and Attendance - Cubic

	Y = 1 if completed at least 1 year of college					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_97	-0.00450 (0.00243)		-0.00170 (0.00229)	-0.00330 (0.00306)		0.000345 (0.00327)
Expectations ²	0.000231*** (0.0000583)		0.000126* (0.0000554)	0.000222** (0.0000705)		0.000100 (0.0000755)
Expectations ³ _1997	-0.00000134*** (0.000000368)		-0.000000824* (0.000000347)	-0.00000132** (0.000000436)		-0.000000702 (0.000000467)
Male (d)		-0.0837*** (0.0150)	-0.0713*** (0.0149)		-0.107*** (0.0191)	-0.0948*** (0.0196)
White (d)		-0.0769** (0.0235)	-0.0672** (0.0232)		-0.106*** (0.0299)	-0.0922** (0.0305)
Black (d)		0.0891*** (0.0242)	0.0647** (0.0240)		0.112*** (0.0289)	0.0895** (0.0300)
Ln(GrossFamIncome)		0.0287** (0.00938)	0.0240** (0.00907)		0.0353** (0.0128)	0.0316* (0.0129)
One parent some college (d)		0.0614*** (0.0186)	0.0486** (0.0183)		0.0689** (0.0227)	0.0567* (0.0230)
Both parents some college (d)		0.101** (0.0352)	0.0864* (0.0351)		0.118** (0.0442)	0.103* (0.0458)
One parent 4yr degree (d)		0.141*** (0.0216)	0.111*** (0.0215)		0.172*** (0.0267)	0.138*** (0.0276)
Both parents 4yr degree (d)		0.0643* (0.0279)	0.0615* (0.0274)		0.137** (0.0467)	0.135** (0.0471)
ASVAB Percentile		0.00497*** (0.000357)	0.00455*** (0.000355)		0.00651*** (0.000502)	0.00616*** (0.000512)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	3450	3450	3450	3450	3450	3450
Adjusted R^2	0.153	0.277	0.304			
Pseudo R^2				0.116	0.241	0.264

Notes:

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.38: 2001 Expectations and Attendance - Cubic

	Y = 1 if completed at least 1 year of college					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_01	-0.0000558 (0.00384)		-0.00152 (0.00351)	0.00289 (0.00592)		0.00249 (0.00672)
Expectations ²	0.0000551 (0.000105)		0.0000583 (0.0000947)	0.0000281 (0.000146)		0.00000498 (0.000164)
Expectations ³ _2001	0.000000133 (0.000000702)		-7.67e-08 (0.000000627)	0.000000235 (0.000000927)		0.000000227 (0.00000103)
Male (d)		-0.0724* (0.0286)	-0.0488 (0.0269)		-0.128** (0.0445)	-0.0954* (0.0459)
White (d)		0.00491 (0.0401)	0.0160 (0.0380)		0.00562 (0.0638)	0.0195 (0.0649)
Black (d)		0.0996* (0.0425)	0.0729 (0.0411)		0.167* (0.0669)	0.135 (0.0701)
Ln(GrossFamIncome)		0.0109 (0.0165)	0.00681 (0.0164)		0.0201 (0.0265)	0.0134 (0.0282)
One parent some college (d)		0.0658 (0.0393)	0.0290 (0.0372)		0.105 (0.0553)	0.0563 (0.0573)
Both parents some college (d)		0.0957 (0.0708)	0.0372 (0.0705)		0.106 (0.0962)	0.0433 (0.0972)
One parent 4yr degree (d)		0.154** (0.0473)	0.0976* (0.0449)		0.250*** (0.0757)	0.165* (0.0783)
Both parents 4yr degree (d)		-0.00211 (0.0710)	-0.0364 (0.0647)		-0.0442 (0.131)	-0.0833 (0.119)
ASVAB Percentile		0.00460*** (0.000775)	0.00383*** (0.000721)		0.00651*** (0.00117)	0.00583*** (0.00116)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	792	792	792	792	792	792
Adjusted R^2	0.321	0.400	0.455			
Pseudo R^2				0.259	0.389	0.446

Notes: 2001 Sub-sample limited to those with either a HS Degree or older than 18 in 2001.

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.39: 1997 Expectations and Completion - Cubic

	Y = 1 if completed a 4 year college degree					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_97	-0.00579** (0.00183)		-0.00207 (0.00175)	-0.00899** (0.00315)		-0.00267 (0.00333)
Expectations ²	0.000166*** (0.0000468)		0.0000476 (0.0000438)	0.000282*** (0.0000695)		0.0000930 (0.0000728)
Expectations ³ _1997	-0.000000715* (0.000000307)		-0.000000155 (0.000000283)	-0.00000150*** (0.000000417)		-0.000000483 (0.000000436)
Male (d)		-0.0447*** (0.0135)	-0.0348** (0.0134)		-0.0570*** (0.0167)	-0.0436** (0.0167)
White (d)		-0.00843 (0.0195)	-0.00480 (0.0194)		-0.00743 (0.0255)	-0.00135 (0.0253)
Black (d)		0.101*** (0.0203)	0.0840*** (0.0203)		0.134*** (0.0311)	0.112*** (0.0308)
Ln(GrossFamIncome)		0.0266*** (0.00768)	0.0244** (0.00759)		0.0448** (0.0145)	0.0414** (0.0141)
One parent some college (d)		0.0246 (0.0171)	0.0169 (0.0169)		0.0302 (0.0203)	0.0210 (0.0198)
Both parents some college (d)		0.0996** (0.0363)	0.0943** (0.0359)		0.109* (0.0425)	0.100* (0.0420)
One parent 4yr degree (d)		0.168*** (0.0219)	0.152*** (0.0221)		0.180*** (0.0252)	0.158*** (0.0248)
Both parents 4yr degree (d)		0.117*** (0.0326)	0.111*** (0.0324)		0.113** (0.0399)	0.104** (0.0391)
ASVAB Percentile		0.00334*** (0.000337)	0.00313*** (0.000335)		0.00367*** (0.000394)	0.00332*** (0.000393)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	3450	3450	3450	3450	3450	3450
Adjusted R^2	0.118	0.307	0.317			
Pseudo R^2				0.111	0.277	0.290

Notes:

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.40: 2001 Expectations and Completion - Cubic

Y = 1 if completed a 4 year college degree						
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_01	0.00431*** (0.000315)		-0.00311 (0.00259)	-0.00823 (0.00501)		-0.00618 (0.00475)
Expectations ²			0.0000746 (0.0000748)	0.000251* (0.000116)		0.000175 (0.000111)
Expectations ³ _2001			-0.000000271 (0.000000518)	-0.00000128 (0.000000707)		-0.000000938 (0.000000677)
Male (d)		-0.0242 (0.0246)	-0.0125 (0.0243)		-0.0293 (0.0288)	-0.0139 (0.0276)
White (d)		-0.0189 (0.0333)	-0.0116 (0.0323)		-0.0264 (0.0409)	-0.0110 (0.0378)
Black (d)		0.00237 (0.0351)	-0.00984 (0.0350)		-0.000983 (0.0458)	-0.00291 (0.0427)
Ln(GrossFamIncome)		0.00484 (0.0137)	0.00197 (0.0140)		0.00456 (0.0180)	-0.00205 (0.0177)
One parent some college (d)		0.0401 (0.0348)	0.0204 (0.0338)		0.0441 (0.0373)	0.0200 (0.0335)
Both parents some college (d)		0.162* (0.0727)	0.132 (0.0738)		0.147 (0.0799)	0.108 (0.0744)
One parent 4yr degree (d)		0.140** (0.0532)	0.111* (0.0530)		0.118* (0.0558)	0.0801 (0.0499)
Both parents 4yr degree (d)		0.217** (0.0824)	0.198* (0.0793)		0.216 (0.112)	0.176 (0.102)
ASVAB Percentile		0.00251*** (0.000730)	0.00214** (0.000703)		0.00214** (0.000687)	0.00164** (0.000632)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	792	792	792	792	758	758
Adjusted R^2	0.164	0.350	0.370			
Pseudo R^2				0.205	0.358	0.393

Notes:

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.41: Expectations Indicators, Attendance, and Completion

	<u>Y = 1 if completed at least 1 year of college</u>	<u>Y = 1 if completed a 4 year college degree</u>	
	<u>1997 Sample</u>	<u>1997 Sample</u>	<u>2001 Sample</u>
	(1)	(2)	(3)
College Expectations_97	0.00397*** (0.000570)	0.00278*** (0.000528)	
Expect=0_97 (d)	-0.0161 (0.0617)	0.109 (0.0803)	
Expect=50_97 (d)	-0.0430 (0.0309)	0.00576 (0.0301)	
Expect=100_97 (d)	-0.0462 (0.0288)	-0.0143 (0.0185)	
College Expectations_01			0.00587*** (0.000805)
Expect=0_01 (d)			0.309 (0.166)
Expect=50_01 (d)			0.0143 (0.0591)
Expect=100_01 (d)			-0.0493* (0.0223)
Demographic Controls	Yes	Yes	Yes
Parent Education Controls	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes
Family Controls	Yes	Yes	Yes
Grades in 8th Grade	Yes	Yes	No
Grades in High School	No	No	Yes
Peer and Teacher Controls	Yes	Yes	Yes
College Enrollment Controls	No	No	Yes
Observations	3403	3366	1815
Pseudo R^2	0.286	0.340	0.487

Notes: Demographic controls include male, white, and black. Parent education include indicators for if one or both of the respondents' parents have some college or a four-year degree. Location controls include rural, 1 of 4 census tracks, and MSA. Family controls include number of members in the household under 18, whether student is only child, and indicators for if one or both parents are missing. Peer controls is the reported percent of peers that the respondent expects to go to college.

-Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

-* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.42: 1997 Expectations and Attendance - Confidence Indicators

	Y = 1 if completed at least 1 year of college					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_97	0.00679*** (0.000379)		0.00334*** (0.000408)	0.00737*** (0.000502)		0.00393*** (0.000553)
Exept=0 in 97 (d)	0.0927* (0.0382)		0.0268 (0.0367)	0.0907 (0.0524)		0.0141 (0.0571)
Exept=50 in 97 (d)	-0.0924*** (0.0247)		-0.0356 (0.0237)	-0.0838** (0.0282)		-0.0321 (0.0299)
Exept=100 in 97 (d)	-0.0686** (0.0221)		-0.0389 (0.0202)	-0.0776** (0.0255)		-0.0505 (0.0284)
Male (d)		-0.0837*** (0.0150)	-0.0712*** (0.0149)		-0.107*** (0.0191)	-0.0948*** (0.0196)
White (d)		-0.0769** (0.0235)	-0.0664** (0.0232)		-0.106*** (0.0299)	-0.0917** (0.0304)
Black (d)		0.0891*** (0.0242)	0.0648** (0.0240)		0.112*** (0.0289)	0.0894** (0.0300)
Lngrossfaminc_1997		0.0287** (0.00938)	0.0241** (0.00909)		0.0353** (0.0128)	0.0313* (0.0128)
One parent some college (d)		0.0614*** (0.0186)	0.0487** (0.0183)		0.0689** (0.0227)	0.0571* (0.0230)
Both parents some college (d)		0.101** (0.0352)	0.0889* (0.0353)		0.118** (0.0442)	0.106* (0.0459)
One parent 4yr degree (d)		0.141*** (0.0216)	0.112*** (0.0214)		0.172*** (0.0267)	0.138*** (0.0276)
Both parents 4yr degree (d)		0.0643* (0.0279)	0.0592* (0.0274)		0.137** (0.0467)	0.135** (0.0471)
ASVAB Percentile		0.00497*** (0.000357)	0.00449*** (0.000357)		0.00651*** (0.000502)	0.00608*** (0.000513)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	3450	3450	3450	3450	3450	3450
Adjusted R^2	0.156	0.277	0.304			
Pseudo R^2				0.118	0.241	0.265

Notes:

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.43: 2001 Expectations and Attendance - Confidence Indicators

	Y = 1 if completed at least 1 year of college					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_01	0.00663*** (0.000655)		0.00337*** (0.000673)	0.00767*** (0.000980)		0.00442*** (0.00113)
Exept=0 in 01 (d)	0.0756 (0.0470)		0.0558 (0.0454)	0.0508 (0.0954)		0.0177 (0.107)
Exept=50 in 01 (d)	-0.149** (0.0462)		-0.0712 (0.0453)	-0.135* (0.0587)		-0.0502 (0.0731)
Exept=100 in 01 (d)	0.102* (0.0488)		0.0928* (0.0431)	0.104 (0.0570)		0.140* (0.0662)
Male (d)		-0.0724* (0.0286)	-0.0464 (0.0270)		-0.128** (0.0445)	-0.0935* (0.0462)
White (d)		0.00491 (0.0401)	0.0176 (0.0379)		0.00562 (0.0638)	0.0179 (0.0648)
Black (d)		0.0996* (0.0425)	0.0727 (0.0410)		0.167* (0.0669)	0.133 (0.0700)
Lngrossfaminc_1997		0.0109 (0.0165)	0.00635 (0.0163)		0.0201 (0.0265)	0.0146 (0.0280)
One parent some college (d)		0.0658 (0.0393)	0.0259 (0.0371)		0.105 (0.0553)	0.0539 (0.0575)
Both parents some college (d)		0.0957 (0.0708)	0.0369 (0.0687)		0.106 (0.0962)	0.0375 (0.0951)
One parent 4yr degree (d)		0.154** (0.0473)	0.0918* (0.0448)		0.250*** (0.0757)	0.152 (0.0792)
Both parents 4yr degree (d)		-0.00211 (0.0710)	-0.0282 (0.0645)		-0.0442 (0.131)	-0.0697 (0.121)
ASVAB Percentile		0.00460*** (0.000775)	0.00388*** (0.000716)		0.00651*** (0.00117)	0.00591*** (0.00116)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	792	792	792	792	792	792
Adjusted R^2	0.322	0.400	0.457			
Pseudo R^2				0.262	0.389	0.449

Notes: Only 2001 High School Sample are only those who've completed ≤ 11 th grade at the time of the survey are used.

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.44: 1997 Expectations and Completion - Confidence Indicators

	Y = 1 if completed a 4 year college degree					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_97	0.00538*** (0.000297)		0.00173*** (0.000293)	0.00735*** (0.000585)		0.00306*** (0.000577)
Exept=0 in 97 (d)	0.173*** (0.0258)		0.0802** (0.0256)	0.333*** (0.0753)		0.163* (0.0766)
Exept=50 in 97 (d)	-0.0721*** (0.0166)		-0.0165 (0.0167)	-0.0323 (0.0303)		0.00395 (0.0323)
Exept=100 in 97 (d)	-0.0202 (0.0230)		0.00748 (0.0199)	-0.0584** (0.0206)		-0.0144 (0.0221)
Male (d)		-0.0447*** (0.0135)	-0.0359** (0.0134)		-0.0570*** (0.0167)	-0.0435** (0.0166)
White (d)		-0.00843 (0.0195)	-0.00481 (0.0194)		-0.00743 (0.0255)	-0.00165 (0.0253)
Black (d)		0.101*** (0.0203)	0.0852*** (0.0203)		0.134*** (0.0311)	0.112*** (0.0308)
Lngrossfaminc_1997		0.0266*** (0.00768)	0.0241** (0.00761)		0.0448** (0.0145)	0.0411** (0.0141)
One parent some college (d)		0.0246 (0.0171)	0.0172 (0.0169)		0.0302 (0.0203)	0.0217 (0.0198)
Both parents some college (d)		0.0996** (0.0363)	0.0938** (0.0359)		0.109* (0.0425)	0.0992* (0.0418)
One parent 4yr degree (d)		0.168*** (0.0219)	0.152*** (0.0221)		0.180*** (0.0252)	0.156*** (0.0247)
Both parents 4yr degree (d)		0.117*** (0.0326)	0.113*** (0.0324)		0.113** (0.0399)	0.104** (0.0390)
ASVAB Percentile		0.00334*** (0.000337)	0.00315*** (0.000337)		0.00367*** (0.000394)	0.00334*** (0.000393)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	3450	3450	3450	3450	3450	3450
Adjusted R^2	0.117	0.307	0.316			
Pseudo R^2				0.113	0.277	0.291

Notes: Full 1997 sub-sample used in the estimation.

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.45: 2001 Expectations and Completion - Confidence Indicators for High School Sample

	Y = 1 if completed a 4 year college degree					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_01	0.00488*** (0.000530)		0.00212*** (0.000521)	0.00630*** (0.000956)		0.00347*** (0.000956)
Exept=0 in 01 (d)	0.125*** (0.0266)		0.0697* (0.0315)	0.290 (0.153)		0.194 (0.137)
Exept=50 in 01 (d)	-0.105*** (0.0275)		-0.0388 (0.0283)	-0.0483 (0.0535)		-0.00279 (0.0610)
Exept=100 in 01 (d)	0.0284 (0.0512)		0.0151 (0.0438)	-0.0270 (0.0331)		-0.0117 (0.0351)
Male (d)		-0.0242 (0.0246)	-0.0123 (0.0244)		-0.0293 (0.0288)	-0.0120 (0.0270)
White (d)		-0.0189 (0.0333)	-0.0107 (0.0324)		-0.0264 (0.0409)	-0.00898 (0.0374)
Black (d)		0.00237 (0.0351)	-0.00797 (0.0351)		-0.000983 (0.0458)	0.00000100 (0.0423)
Lngrossfaminc_1997		0.00484 (0.0137)	0.00220 (0.0141)		0.00456 (0.0180)	-0.00315 (0.0170)
One parent some college (d)		0.0401 (0.0348)	0.0207 (0.0339)		0.0441 (0.0373)	0.0205 (0.0333)
Both parents some college (d)		0.162* (0.0727)	0.133 (0.0736)		0.147 (0.0799)	0.104 (0.0732)
One parent 4yr degree (d)		0.140** (0.0532)	0.112* (0.0532)		0.118* (0.0558)	0.0798 (0.0496)
Both parents 4yr degree (d)		0.217** (0.0824)	0.201* (0.0796)		0.216 (0.112)	0.176 (0.102)
ASVAB Percentile		0.00251*** (0.000730)	0.00216** (0.000703)		0.00214** (0.000687)	0.00161** (0.000621)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	792	792	792	792	758	758
Adjusted R^2	0.181	0.350	0.367			
Pseudo R^2				0.206	0.358	0.393

Notes: 2001 High School Sample used are only those who've completed ≤ 11 th grade at the time of the survey. Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.46: 2001 Expectations and Completion w/ Enrollment and Confidence Indicators

	Y = 1 if completed a 4 year college degree					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_01	0.00549*** (0.000367)		0.00220*** (0.000321)	0.00893*** (0.000999)		0.00496*** (0.000901)
Exept=0 in 01 (d)	0.140*** (0.0189)		0.0688*** (0.0199)	0.379** (0.136)		0.187 (0.118)
Exept=50 in 01 (d)	-0.110*** (0.0201)		-0.0355 (0.0202)	-0.0446 (0.0574)		-0.0138 (0.0623)
Exept=100 in 01 (d)	0.126*** (0.0319)		0.0532* (0.0263)	0.0374 (0.0301)		0.0170 (0.0325)
Male (d)		-0.0360* (0.0161)	-0.0235 (0.0158)		-0.0610* (0.0255)	-0.0401 (0.0246)
White (d)		0.00569 (0.0223)	0.0242 (0.0221)		0.00336 (0.0358)	0.0305 (0.0344)
Black (d)		0.0159 (0.0247)	0.00769 (0.0243)		0.0267 (0.0423)	0.00658 (0.0396)
Lngrossfaminc_1997		0.00710 (0.00962)	0.00123 (0.00960)		0.00879 (0.0175)	-0.00238 (0.0171)
One parent some college (d)		0.0192 (0.0197)	0.00774 (0.0193)		0.0207 (0.0291)	0.00794 (0.0279)
Both parents some college (d)		0.100* (0.0413)	0.0782 (0.0405)		0.152* (0.0601)	0.114* (0.0578)
One parent 4yr degree (d)		0.126*** (0.0278)	0.102*** (0.0275)		0.170*** (0.0391)	0.130*** (0.0376)
Both parents 4yr degree (d)		0.117** (0.0402)	0.111** (0.0390)		0.170* (0.0681)	0.136* (0.0634)
ASVAB Percentile		0.00140*** (0.000421)	0.00104* (0.000411)		0.00174** (0.000603)	0.00107 (0.000575)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Enrollment Controls	No	Yes	Yes	No	Yes	Yes
Observations	1902	1902	1902	1902	1850	1850
Adjusted R^2	0.259	0.465	0.490			
Pseudo R^2				0.254	0.425	0.467

Notes:

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.47: Expectations and Attendance with Parent Expectations

	Y = 1 if completed a 4 year college degree			
	OLS	OLS	Probit	Probit
College Expectations_97	0.00321*** (0.000294)	0.00229*** (0.000329)	0.00382*** (0.000377)	0.00280*** (0.000412)
ParentExpectations_1997		0.00224*** (0.000332)		0.00266*** (0.000408)
Male (d)	-0.0591*** (0.0160)	-0.0592*** (0.0159)	-0.0779*** (0.0214)	-0.0790*** (0.0215)
White (d)	-0.0539* (0.0254)	-0.0367 (0.0254)	-0.0753* (0.0335)	-0.0544 (0.0339)
Black (d)	0.0940*** (0.0263)	0.0927*** (0.0261)	0.129*** (0.0322)	0.128*** (0.0323)
Lngrossfaminc_1997	0.0199* (0.00941)	0.0157 (0.00934)	0.0250 (0.0134)	0.0194 (0.0133)
One parent some college (d)	0.0413* (0.0195)	0.0314 (0.0193)	0.0499* (0.0246)	0.0363 (0.0248)
Both parents some college (d)	0.0741* (0.0363)	0.0714* (0.0360)	0.0915 (0.0478)	0.0918 (0.0480)
One parent 4yr degree (d)	0.0840*** (0.0226)	0.0758*** (0.0225)	0.104*** (0.0293)	0.0950** (0.0297)
Both parents 4yr degree (d)	0.0649* (0.0294)	0.0592* (0.0291)	0.138** (0.0486)	0.132** (0.0494)
ASVAB Percentile	0.00455*** (0.000381)	0.00419*** (0.000381)	0.00612*** (0.000545)	0.00571*** (0.000550)
Location Controls	Yes	Yes	Yes	Yes
Family Controls	Yes	Yes	Yes	Yes
Grades in 8th Grade	Yes	Yes	Yes	Yes
Teacher and Peer Controls	Yes	Yes	Yes	Yes
Observations	2938	2938	2938	2938
Adjusted R^2	0.310	0.322		
Pseudo R^2			0.272	0.282

Notes:

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.48: Parent Expectations and College Achievement

	Y = 1 if completed at least 1 year of college	Y = 1 if completed a 4 year college degree		
	Full 1997 Sample	Full 1997 Sample	1997 and 2001 Respondents	
	(1)	(2)	(3)	(4)
College Expectations_97	0.00310*** (0.000424)	0.00182*** (0.000379)	0.00215* (0.000898)	
College Expectations_01				0.00365*** (0.000528)
Parent Expectations_97	0.00276*** (0.000419)	0.00208*** (0.000360)	0.00306*** (0.000916)	0.000912 (0.000629)
Demographic Controls	Yes	Yes	Yes	Yes
Parent Education Controls	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes
Family Controls	Yes	Yes	Yes	Yes
Grades in 8th Grade	Yes	Yes	Yes	No
Grades in High School	No	No	No	Yes
Peer Controls	Yes	Yes	Yes	Yes
College Enrollment Controls	No	No	No	Yes
Observations	2903	2870	569	542
Pseudo R^2	0.304	0.353	0.400	0.601

Notes: Demographic controls include male, white, and black. Parent education include indicators for if one or both of the respondents' parents have some college or a four-year degree. Location controls include rural, 1 of 4 census tracks, and MSA. Family controls include number of members in the household under 18, whether student is only child, and indicators for if one or both parents are missing. Peer controls is the reported percent of peers that the respondent expects to go to college.

-Marginal effects; Robust standard errors in parentheses

—* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.49: Expectations and Completion with Parent Expectations

	Y = 1 if completed a 4 year college degree							
	OLS	OLS	Probit	Probit	OLS	OLS	Probit	Probit
College Expectations_97	0.00166*** (0.000224)	0.00103*** (0.000246)	0.00270*** (0.000376)	0.00177*** (0.000416)				
College Expectations_01					0.00164*** (0.000336)	0.00159*** (0.000338)	0.00398*** (0.000634)	0.00383*** (0.000640)
ParentExpectations_1997		0.00154*** (0.000260)		0.00223*** (0.000403)		0.000537 (0.000439)		0.00122 (0.000815)
Male (d)	-0.0248 (0.0145)	-0.0249 (0.0145)	-0.0283 (0.0182)	-0.0307 (0.0181)	-0.0100 (0.0257)	-0.00843 (0.0257)	-0.0572 (0.0395)	-0.0539 (0.0399)
White (d)	0.000341 (0.0210)	0.0122 (0.0211)	0.00525 (0.0279)	0.0177 (0.0279)	0.0586 (0.0369)	0.0622 (0.0378)	0.0881 (0.0576)	0.0903 (0.0578)
Black (d)	0.112*** (0.0222)	0.111*** (0.0221)	0.151*** (0.0347)	0.146*** (0.0346)	0.0335 (0.0411)	0.0293 (0.0416)	0.0207 (0.0717)	0.0113 (0.0721)
Lngrossfaminc_1997	0.0223** (0.00792)	0.0195* (0.00781)	0.0389** (0.0146)	0.0350* (0.0142)	0.00327 (0.0106)	0.00217 (0.0106)	0.0267 (0.0305)	0.0230 (0.0299)
One parent some college (d)	0.0153 (0.0182)	0.00851 (0.0181)	0.0175 (0.0214)	0.0109 (0.0211)	-0.0244 (0.0309)	-0.0255 (0.0309)	-0.0186 (0.0470)	-0.0181 (0.0471)
Both parents some college (d)	0.0905* (0.0371)	0.0886* (0.0370)	0.0937* (0.0428)	0.0915* (0.0426)	0.0946 (0.0683)	0.0886 (0.0685)	0.142 (0.109)	0.127 (0.105)
One parent 4yr degree (d)	0.152*** (0.0237)	0.146*** (0.0235)	0.159*** (0.0266)	0.152*** (0.0264)	0.140** (0.0466)	0.137** (0.0469)	0.194** (0.0684)	0.189** (0.0681)
Both parents 4yr degree (d)	0.110** (0.0349)	0.106** (0.0348)	0.0993* (0.0420)	0.0953* (0.0418)	-0.0355 (0.0671)	-0.0344 (0.0671)	-0.0311 (0.0738)	-0.0317 (0.0728)
ASVAB Percentile	0.00315*** (0.000362)	0.00290*** (0.000362)	0.00339*** (0.000426)	0.00302*** (0.000423)	-0.000175 (0.000703)	-0.000283 (0.000712)	0.000982 (0.000974)	0.000668 (0.00100)
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Family Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grades in 8th Grade	Yes	Yes	Yes	Yes	No	No	No	No
Grades in HS	No	No	No	No	Yes	Yes	No	No
Teacher and Peer Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
College Enrollment Controls	No	No	No	No	Yes	Yes	Yes	Yes
Observations	2938	2938	2938	2938	611	611	611	611
Adjusted R^2	0.321	0.328			0.574	0.574		
Pseudo R^2			0.297	0.306			0.555	0.557

Notes: Estimates from the 700 overlapping students are included in the last 4 columns

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.50: 1997 Expectations and Attendance - Enrollment Interaction

	Y = 1 if completed at least 1 year of college					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_97	0.00328*** (0.000741)		0.00146* (0.000720)	0.00375*** (0.000809)		0.00177* (0.000896)
Enroll in HS X Expectations_1997	0.00274*** (0.000700)		0.00151* (0.000685)	0.00281*** (0.000747)		0.00173* (0.000842)
Male (d)		-0.0837*** (0.0150)	-0.0699*** (0.0148)		-0.107*** (0.0191)	-0.0931*** (0.0195)
White (d)		-0.0769** (0.0235)	-0.0644** (0.0231)		-0.106*** (0.0299)	-0.0898** (0.0304)
Black (d)		0.0891*** (0.0242)	0.0627** (0.0239)		0.112*** (0.0289)	0.0862** (0.0299)
Ln(GrossFamIncome)		0.0287** (0.00938)	0.0224* (0.00916)		0.0353** (0.0128)	0.0294* (0.0130)
One parent some college (d)		0.0614*** (0.0186)	0.0478** (0.0183)		0.0689** (0.0227)	0.0556* (0.0230)
Both parents some college (d)		0.101** (0.0352)	0.0858* (0.0351)		0.118** (0.0442)	0.103* (0.0458)
One parent 4yr degree (d)		0.141*** (0.0216)	0.113*** (0.0214)		0.172*** (0.0267)	0.139*** (0.0276)
Both parents 4yr degree (d)		0.0643* (0.0279)	0.0583* (0.0274)		0.137** (0.0467)	0.132** (0.0473)
ASVAB Percentile		0.00497*** (0.000357)	0.00458*** (0.000353)		0.00651*** (0.000502)	0.00619*** (0.000510)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	3450	3450	3450	3450	3450	3450
Adjusted R^2	0.153	0.277	0.304			
Pseudo R^2				0.117	0.241	0.265

Notes:

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.51: 2001 Expectations and Attendance - Enrollment Interaction

	Y = 1 if completed at least 1 year of college					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_01	0.00170* (0.000662)		0.00143* (0.000672)	0.00268** (0.00103)		0.00447*** (0.00136)
Enroll in HS X Expectations_2001	0.00558*** (0.000605)		0.00320*** (0.000855)	0.00605*** (0.000848)		0.00128 (0.00160)
Male (d)		-0.0724* (0.0286)	-0.0530* (0.0268)		-0.128** (0.0445)	-0.100* (0.0452)
White (d)		0.00491 (0.0401)	0.0165 (0.0378)		0.00562 (0.0638)	0.0193 (0.0642)
Black (d)		0.0996* (0.0425)	0.0796 (0.0406)		0.167* (0.0669)	0.143* (0.0694)
Ln(GrossFamIncome)		0.0109 (0.0165)	0.00732 (0.0160)		0.0201 (0.0265)	0.0150 (0.0278)
One parent some college (d)		0.0658 (0.0393)	0.0341 (0.0369)		0.105 (0.0553)	0.0612 (0.0573)
Both parents some college (d)		0.0957 (0.0708)	0.0469 (0.0690)		0.106 (0.0962)	0.0513 (0.0978)
One parent 4yr degree (d)		0.154** (0.0473)	0.0997* (0.0452)		0.250*** (0.0757)	0.167* (0.0782)
Both parents 4yr degree (d)		-0.00211 (0.0710)	-0.0374 (0.0657)		-0.0442 (0.131)	-0.0806 (0.121)
ASVAB Percentile		0.00460*** (0.000775)	0.00371*** (0.000722)		0.00651*** (0.00117)	0.00579*** (0.00115)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	792	792	792	792	792	792
Adjusted R^2	0.372	0.400	0.458			
Pseudo R^2				0.305	0.389	0.444

Notes:

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.52: 1997 Expectations and Completion - Enrollment Interaction

	Y = 1 if completed a 4 year college degree					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_97	0.00287*** (0.000589)		0.00109 (0.000572)	0.00403*** (0.000779)		0.00208* (0.000848)
Enroll in HS X Expectations_1997	0.00182** (0.000564)		0.000389 (0.000547)	0.00172* (0.000701)		0.000228 (0.000779)
Male (d)		-0.0447*** (0.0135)	-0.0377** (0.0134)		-0.0570*** (0.0167)	-0.0445** (0.0166)
White (d)		-0.00843 (0.0195)	-0.00256 (0.0194)		-0.00743 (0.0255)	-0.000296 (0.0253)
Black (d)		0.101*** (0.0203)	0.0876*** (0.0203)		0.134*** (0.0311)	0.114*** (0.0308)
Ln(GrossFamIncome)		0.0266*** (0.00768)	0.0239** (0.00760)		0.0448** (0.0145)	0.0412** (0.0141)
One parent some college (d)		0.0246 (0.0171)	0.0177 (0.0170)		0.0302 (0.0203)	0.0216 (0.0198)
Both parents some college (d)		0.0996** (0.0363)	0.0926* (0.0361)		0.109* (0.0425)	0.0993* (0.0419)
One parent 4yr degree (d)		0.168*** (0.0219)	0.153*** (0.0221)		0.180*** (0.0252)	0.160*** (0.0249)
Both parents 4yr degree (d)		0.117*** (0.0326)	0.114*** (0.0325)		0.113** (0.0399)	0.105** (0.0392)
ASVAB Percentile		0.00334*** (0.000337)	0.00314*** (0.000335)		0.00367*** (0.000394)	0.00334*** (0.000391)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	3450	3450	3450	3450	3450	3450
Adjusted R^2	0.109	0.307	0.315			
Pseudo R^2				0.107	0.277	0.289

Notes:

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.53: 2001 Expectations and Completion - Enrollment Interaction

	Y = 1 if completed a 4 year college degree					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_01	0.00105* (0.000439)		0.000242 (0.000467)	0.00176* (0.000871)		0.00162 (0.00104)
Enroll in HS X Expectations_2001	0.00346*** (0.000445)		0.00221*** (0.000619)	0.00301*** (0.000701)		0.00114 (0.00126)
Male (d)		-0.0242 (0.0246)	-0.0152 (0.0243)		-0.0293 (0.0288)	-0.0146 (0.0277)
White (d)		-0.0189 (0.0333)	-0.0116 (0.0325)		-0.0264 (0.0409)	-0.0132 (0.0384)
Black (d)		0.00237 (0.0351)	-0.00557 (0.0350)		-0.000983 (0.0458)	-0.00356 (0.0435)
Ln(GrossFamIncome)		0.00484 (0.0137)	0.00236 (0.0139)		0.00456 (0.0180)	-0.000872 (0.0174)
One parent some college (d)		0.0401 (0.0348)	0.0239 (0.0337)		0.0441 (0.0373)	0.0209 (0.0341)
Both parents some college (d)		0.162* (0.0727)	0.138 (0.0732)		0.147 (0.0799)	0.115 (0.0765)
One parent 4yr degree (d)		0.140** (0.0532)	0.112* (0.0524)		0.118* (0.0558)	0.0817 (0.0503)
Both parents 4yr degree (d)		0.217** (0.0824)	0.198* (0.0797)		0.216 (0.112)	0.182 (0.104)
ASVAB Percentile		0.00251*** (0.000730)	0.00204** (0.000701)		0.00214** (0.000687)	0.00163* (0.000637)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	792	792	792	792	758	758
Adjusted R^2	0.203	0.350	0.372			
Pseudo R^2				0.226	0.358	0.389

Notes:

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.54: 1997 Expectations and Attendance - Age Interaction

	Y = 1 if completed at least 1 year of college					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_97	-0.000632 (0.00330)		-0.00391 (0.00296)	-0.000949 (0.00370)		-0.00522 (0.00406)
Age x Expectations	0.000405* (0.000200)		0.000416* (0.000179)	0.000458* (0.000225)		0.000530* (0.000246)
Male (d)		-0.0837*** (0.0150)	-0.0693*** (0.0148)		-0.107*** (0.0191)	-0.0919*** (0.0195)
White (d)		-0.0769** (0.0235)	-0.0660** (0.0231)		-0.106*** (0.0299)	-0.0920** (0.0303)
Black (d)		0.0891*** (0.0242)	0.0620** (0.0239)		0.112*** (0.0289)	0.0842** (0.0300)
Ln(GrossFamIncome)		0.0287** (0.00938)	0.0246** (0.00911)		0.0353** (0.0128)	0.0317* (0.0129)
One parent some college (d)		0.0614*** (0.0186)	0.0483** (0.0183)		0.0689** (0.0227)	0.0568* (0.0230)
Both parents some college (d)		0.101** (0.0352)	0.0878* (0.0351)		0.118** (0.0442)	0.105* (0.0457)
One parent 4yr degree (d)		0.141*** (0.0216)	0.113*** (0.0215)		0.172*** (0.0267)	0.139*** (0.0276)
Both parents 4yr degree (d)		0.0643* (0.0279)	0.0594* (0.0274)		0.137** (0.0467)	0.134** (0.0472)
ASVAB Percentile		0.00497*** (0.000357)	0.00458*** (0.000354)		0.00651*** (0.000502)	0.00617*** (0.000511)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	3450	3450	3450	3450	3450	3450
Adjusted R^2	0.150	0.277	0.304			
Pseudo R^2				0.115	0.241	0.265

Notes:

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.55: 2001 Expectations and Attendance - Age Interaction

	Y = 1 if completed at least 1 year of college					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_01	0.0425*** (0.00382)		0.0221*** (0.00467)	0.0484*** (0.00564)		0.0236** (0.00723)
Age x Expectations	-0.00204*** (0.000220)		-0.00103*** (0.000256)	-0.00229*** (0.000322)		-0.00102* (0.000402)
Male (d)		-0.0724* (0.0286)	-0.0468 (0.0268)		-0.128** (0.0445)	-0.0944* (0.0461)
White (d)		0.00491 (0.0401)	0.00953 (0.0380)		0.00562 (0.0638)	0.0174 (0.0655)
Black (d)		0.0996* (0.0425)	0.0748 (0.0405)		0.167* (0.0669)	0.145* (0.0704)
Ln(GrossFamIncome)		0.0109 (0.0165)	0.0108 (0.0161)		0.0201 (0.0265)	0.0196 (0.0285)
One parent some college (d)		0.0658 (0.0393)	0.0286 (0.0372)		0.105 (0.0553)	0.0559 (0.0577)
Both parents some college (d)		0.0957 (0.0708)	0.0531 (0.0676)		0.106 (0.0962)	0.0612 (0.0980)
One parent 4yr degree (d)		0.154** (0.0473)	0.104* (0.0454)		0.250*** (0.0757)	0.170* (0.0798)
Both parents 4yr degree (d)		-0.00211 (0.0710)	-0.0404 (0.0658)		-0.0442 (0.131)	-0.0907 (0.121)
ASVAB Percentile		0.00460*** (0.000775)	0.00366*** (0.000720)		0.00651*** (0.00117)	0.00568*** (0.00117)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	792	792	792	792	792	792
Adjusted R^2	0.369	0.400	0.462			
Pseudo R^2				0.305	0.389	0.450

Notes:

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.56: 1997 Expectations and Completion - Age Interaction

	Y = 1 if completed a 4 year college degree					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_97	0.00492 (0.00341)		0.00173 (0.00295)	0.00594* (0.00300)		0.00272 (0.00316)
Age x Expectations	-0.0000138 (0.000207)		-0.0000157 (0.000179)	-0.00000998 (0.000182)		-0.0000251 (0.000191)
Male (d)		-0.0447*** (0.0135)	-0.0375** (0.0134)		-0.0570*** (0.0167)	-0.0443** (0.0166)
White (d)		-0.00843 (0.0195)	-0.00299 (0.0194)		-0.00743 (0.0255)	-0.000591 (0.0253)
Black (d)		0.101*** (0.0203)	0.0878*** (0.0203)		0.134*** (0.0311)	0.114*** (0.0308)
Ln(GrossFamIncome)		0.0266*** (0.00768)	0.0244** (0.00761)		0.0448** (0.0145)	0.0416** (0.0141)
One parent some college (d)		0.0246 (0.0171)	0.0175 (0.0170)		0.0302 (0.0203)	0.0213 (0.0198)
Both parents some college (d)		0.0996** (0.0363)	0.0930** (0.0361)		0.109* (0.0425)	0.0994* (0.0419)
One parent 4yr degree (d)		0.168*** (0.0219)	0.154*** (0.0221)		0.180*** (0.0252)	0.160*** (0.0249)
Both parents 4yr degree (d)		0.117*** (0.0326)	0.114*** (0.0325)		0.113** (0.0399)	0.105** (0.0392)
ASVAB Percentile		0.00334*** (0.000337)	0.00314*** (0.000335)		0.00367*** (0.000394)	0.00334*** (0.000390)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	3450	3450	3450	3450	3450	3450
Adjusted R^2	0.107	0.307	0.315			
Pseudo R^2				0.105	0.277	0.289

Notes:

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1
 * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.57: 2001 Expectations and Completion - Age Interaction

	Y = 1 if completed a 4 year college degree					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_01	0.0237*** (0.00321)		0.00745* (0.00373)	0.0195*** (0.00336)		0.00554 (0.00414)
Age x Expectations	-0.00111*** (0.000179)		-0.000319 (0.000201)	-0.000846*** (0.000194)		-0.000173 (0.000229)
Male (d)		-0.0242 (0.0246)	-0.0126 (0.0244)		-0.0293 (0.0288)	-0.0121 (0.0275)
White (d)		-0.0189 (0.0333)	-0.0170 (0.0327)		-0.0264 (0.0409)	-0.0161 (0.0385)
Black (d)		0.00237 (0.0351)	-0.0102 (0.0351)		-0.000983 (0.0458)	-0.00678 (0.0428)
Ln(GrossFamIncome)		0.00484 (0.0137)	0.00478 (0.0140)		0.00456 (0.0180)	0.000555 (0.0174)
One parent some college (d)		0.0401 (0.0348)	0.0239 (0.0338)		0.0441 (0.0373)	0.0198 (0.0334)
Both parents some college (d)		0.162* (0.0727)	0.140 (0.0737)		0.147 (0.0799)	0.114 (0.0754)
One parent 4yr degree (d)		0.140** (0.0532)	0.117* (0.0530)		0.118* (0.0558)	0.0811 (0.0499)
Both parents 4yr degree (d)		0.217** (0.0824)	0.200* (0.0802)		0.216 (0.112)	0.178 (0.103)
ASVAB Percentile		0.00251*** (0.000730)	0.00209** (0.000709)		0.00214** (0.000687)	0.00161* (0.000631)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	792	792	792	792	758	758
Adjusted R^2	0.192	0.350	0.367			
Pseudo R^2				0.215	0.358	0.389

Notes:

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1
 * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.58: 1997 Expectations and Attendance - Highest Grade Completed Interaction

	Y = 1 if completed at least 1 year of college					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_97	-0.00498*** (0.00103)		-0.00272** (0.000943)	-0.00569*** (0.00123)		-0.00336* (0.00131)
Highest Grade Comp x Expectations	0.00118*** (0.000106)		0.000624*** (0.000101)	0.00132*** (0.000129)		0.000752*** (0.000139)
Male (d)		-0.0837*** (0.0150)	-0.0615*** (0.0148)		-0.107*** (0.0191)	-0.0825*** (0.0196)
White (d)		-0.0769** (0.0235)	-0.0628** (0.0231)		-0.106*** (0.0299)	-0.0883** (0.0304)
Black (d)		0.0891*** (0.0242)	0.0586* (0.0238)		0.112*** (0.0289)	0.0806** (0.0300)
Ln(GrossFamIncome)		0.0287** (0.00938)	0.0222* (0.00900)		0.0353** (0.0128)	0.0287* (0.0127)
One parent some college (d)		0.0614*** (0.0186)	0.0456* (0.0181)		0.0689** (0.0227)	0.0537* (0.0231)
Both parents some college (d)		0.101** (0.0352)	0.0911** (0.0351)		0.118** (0.0442)	0.109* (0.0459)
One parent 4yr degree (d)		0.141*** (0.0216)	0.116*** (0.0213)		0.172*** (0.0267)	0.142*** (0.0275)
Both parents 4yr degree (d)		0.0643* (0.0279)	0.0561* (0.0275)		0.137** (0.0467)	0.132** (0.0478)
ASVAB Percentile		0.00497*** (0.000357)	0.00428*** (0.000356)		0.00651*** (0.000502)	0.00582*** (0.000514)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	3450	3450	3450	3450	3450	3450
Adjusted R^2	0.181	0.277	0.311			
Pseudo R^2				0.140	0.241	0.270

Notes:

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.59: 2001 Expectations and Attendance - Highest Grade Completed Interaction

	Y = 1 if completed at least 1 year of college					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_01	-0.00500 (0.00687)		-0.000322 (0.00417)	-0.00241 (0.00655)		0.00312 (0.00375)
Highest Grade Comp x Expectations	0.00113 (0.000640)		0.000381 (0.000401)	0.00102 (0.000608)		0.000225 (0.000355)
Male (d)		-0.0724* (0.0286)	-0.0476 (0.0270)		-0.128** (0.0445)	-0.0942* (0.0457)
White (d)		0.00491 (0.0401)	0.0107 (0.0381)		0.00562 (0.0638)	0.0157 (0.0650)
Black (d)		0.0996* (0.0425)	0.0667 (0.0411)		0.167* (0.0669)	0.133 (0.0700)
Ln(GrossFamIncome)		0.0109 (0.0165)	0.0105 (0.0162)		0.0201 (0.0265)	0.0165 (0.0279)
One parent some college (d)		0.0658 (0.0393)	0.0382 (0.0373)		0.105 (0.0553)	0.0626 (0.0571)
Both parents some college (d)		0.0957 (0.0708)	0.0438 (0.0705)		0.106 (0.0962)	0.0482 (0.0968)
One parent 4yr degree (d)		0.154** (0.0473)	0.106* (0.0456)		0.250*** (0.0757)	0.167* (0.0782)
Both parents 4yr degree (d)		-0.00211 (0.0710)	-0.0293 (0.0656)		-0.0442 (0.131)	-0.0803 (0.119)
ASVAB Percentile		0.00460*** (0.000775)	0.00370*** (0.000756)		0.00651*** (0.00117)	0.00572*** (0.00116)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	792	792	792	792	792	792
Adjusted R^2	0.325	0.400	0.451			
Pseudo R^2				0.268	0.389	0.444

Notes:

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.60: 1997 Expectations and Completion - Highest Grade Completed Interaction

	Y = 1 if completed a 4 year college degree					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_97	-0.00375*** (0.000974)		-0.00158 (0.000890)	-0.00176 (0.00100)		-0.000864 (0.00109)
Highest Grade Comp x Expectations	0.000900*** (0.000103)		0.000337*** (0.0000976)	0.000796*** (0.000100)		0.000342** (0.000111)
Male (d)		-0.0447*** (0.0135)	-0.0333* (0.0135)		-0.0570*** (0.0167)	-0.0406* (0.0167)
White (d)		-0.00843 (0.0195)	-0.00121 (0.0194)		-0.00743 (0.0255)	0.000239 (0.0253)
Black (d)		0.101*** (0.0203)	0.0851*** (0.0203)		0.134*** (0.0311)	0.109*** (0.0308)
Ln(GrossFamIncome)		0.0266*** (0.00768)	0.0233** (0.00753)		0.0448** (0.0145)	0.0402** (0.0139)
One parent some college (d)		0.0246 (0.0171)	0.0166 (0.0169)		0.0302 (0.0203)	0.0206 (0.0197)
Both parents some college (d)		0.0996** (0.0363)	0.0951** (0.0362)		0.109* (0.0425)	0.103* (0.0423)
One parent 4yr degree (d)		0.168*** (0.0219)	0.155*** (0.0220)		0.180*** (0.0252)	0.163*** (0.0249)
Both parents 4yr degree (d)		0.117*** (0.0326)	0.112*** (0.0325)		0.113** (0.0399)	0.104** (0.0392)
ASVAB Percentile		0.00334*** (0.000337)	0.00298*** (0.000338)		0.00367*** (0.000394)	0.00318*** (0.000393)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	3450	3450	3450	3450	3450	3450
Adjusted R^2	0.129	0.307	0.318			
Pseudo R^2				0.121	0.277	0.292

Notes:

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.61: 2001 Expectations and Completion - Highest Grade Completed Interaction

	Y = 1 if completed a 4 year college degree					
	OLS	OLS	OLS	Probit	Probit	Probit
College Expectations_01	-0.000325 (0.00461)		0.00296 (0.00282)	0.00244 (0.00279)		0.00503*** (0.00149)
Highest Grade Comp x Expectations	0.000436 (0.000432)		-0.000121 (0.000272)	0.000231 (0.000255)		-0.000249 (0.000145)
Male (d)		-0.0242 (0.0246)	-0.0152 (0.0244)		-0.0293 (0.0288)	-0.0165 (0.0272)
White (d)		-0.0189 (0.0333)	-0.0183 (0.0328)		-0.0264 (0.0409)	-0.0179 (0.0383)
Black (d)		0.00237 (0.0351)	-0.00977 (0.0353)		-0.000983 (0.0458)	-0.00331 (0.0427)
Ln(GrossFamIncome)		0.00484 (0.0137)	0.00488 (0.0139)		0.00456 (0.0180)	0.0000674 (0.0169)
One parent some college (d)		0.0401 (0.0348)	0.0271 (0.0339)		0.0441 (0.0373)	0.0193 (0.0328)
Both parents some college (d)		0.162* (0.0727)	0.139 (0.0737)		0.147 (0.0799)	0.114 (0.0744)
One parent 4yr degree (d)		0.140** (0.0532)	0.120* (0.0529)		0.118* (0.0558)	0.0834 (0.0496)
Both parents 4yr degree (d)		0.217** (0.0824)	0.204* (0.0804)		0.216 (0.112)	0.180 (0.104)
ASVAB Percentile		0.00251*** (0.000730)	0.00221** (0.000722)		0.00214** (0.000687)	0.00171** (0.000625)
Location Controls	No	Yes	Yes	No	Yes	Yes
Family Controls	No	Yes	Yes	No	Yes	Yes
Grades in 8th Grade	No	Yes	Yes	No	Yes	Yes
Teacher and Peer Controls	No	Yes	Yes	No	Yes	Yes
Observations	792	792	792	792	758	758
Adjusted R^2	0.168	0.350	0.365			
Pseudo R^2				0.196	0.358	0.391

Notes:

Marginal effects; Robust standard errors in parentheses; (d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

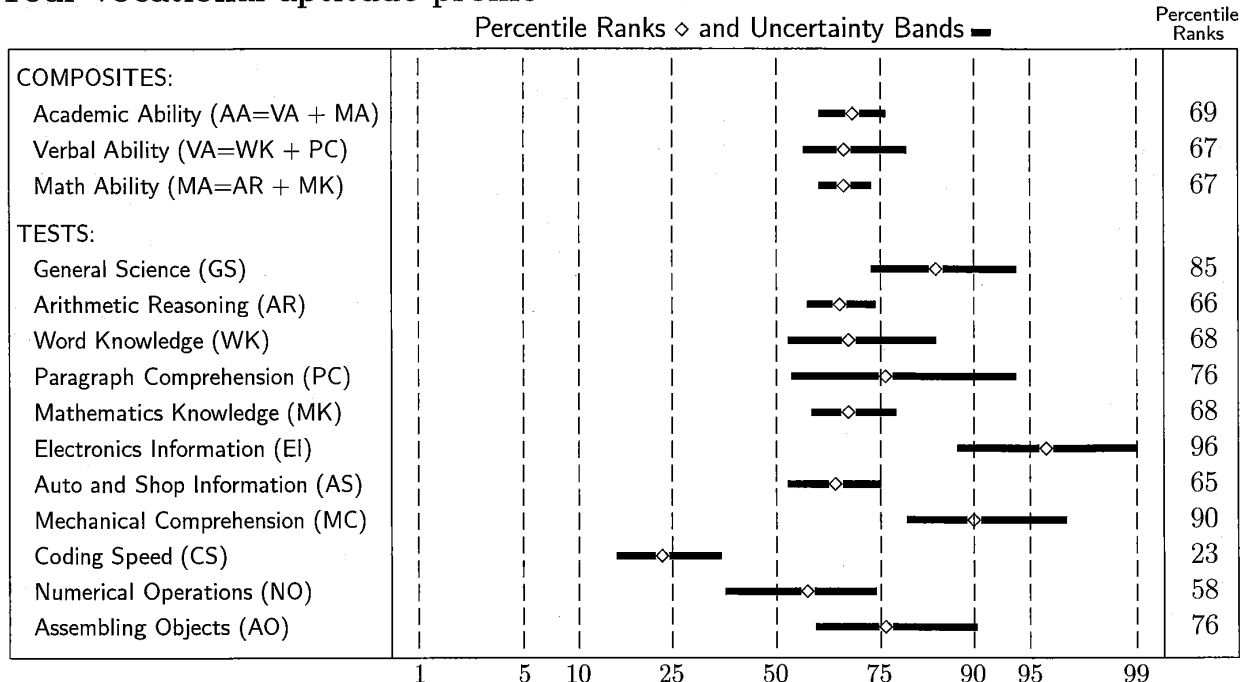
The Profile of American Youth 1997

NORC at the University of Chicago

ID: S999999
Name: Joseph M. Respondent
Address: 84 Rolling Hills Road
Upstate NY 11111E

Comparison Group: 18 to 23 year-olds

Your vocational aptitude profile



Your vocational interest code

R	S	E	C	A	I	R → Realistic	A → Artistic	E → Enterprising
						I → Investigative	S → Social	C → Conventional
1st	2nd	3rd	4th	5th	6th			

How to read your aptitude scores and interest code

Your scores on the Computer Adaptive Test-Armed Services Vocational Aptitude Battery (CAT-ASVAB) and the Interest Finder (IF) are shown in the charts above. Your aptitude scores are reported as percentile ranks and bands. The ranks show how you scored in comparison to a national sample of individuals who were about the same age as you when they took the tests as part of the Profile of American Youth 1997. The ranks indicate the percentage of individuals in your comparison group who scored below you in each area. The meaning of the bands is explained on the back of this form. The comparison group for your scores is shown in the upper right-hand corner of this form.

The results from your answers to the Interest Finder are presented as a six-letter code. The letters are numbered according to their position in your code. Each letter refers to one of six vocational areas—Realistic, Investigative, Artistic, Social, Enterprising, and Conventional—shown to the right of your code. The order of the letters indicates the relative strength of your interests in those areas, as measured by the number of likes and dislikes you reported for a wide variety of activities and jobs in each area. You reported most interest in the first area, second most interest in the second area, and so on.

To learn more about your aptitude scores and interest code please refer to the back of this form.

More About How to Read Your Aptitude Scores and Interest Code

How to Interpret the Percentile Ranks and Bands

Your percentile ranks show how you scored relative to other youth in the nation who were in the same age group as you when they took the aptitude tests. These ranks, like any other test scores, are only approximate. If you were to take the tests again, your scores would probably differ somewhat. The percentile bands show the range of scores you might receive. Each band has a two-thirds chance of including your true score—the score you would receive if your abilities could be perfectly measured.

As you look at your aptitude profile, you will see that the bands differ in size. Some bands cover a wider range of ranks than other bands. The size of each band indicates how accurately the test measures your ability. Tests with shorter bands provide more accurate scores. More accurate scores are likely to be closer in value to your true score than less accurate scores.

How to Interpret Your Aptitude Profile

Your aptitude profile shows how you performed in three general and eleven specific aptitude areas. The position of the bands in the graph indicates the relative strength of your performance in those areas. If the bands from two areas overlap, then you performed at roughly the same level in both areas. If one band is in a higher range than another, then you performed better in that area. To identify your strongest areas of performance, look for areas with bands toward the right-hand side of the graph.

How to Interpret Your Interest Code

On the Interest Finder, you were asked to report your likes and dislikes for a wide range of activities and jobs in six different vocational areas. Your interest code was assigned by counting the number of likes you reported in each area and ordering the areas from highest to lowest "number of likes". It reflects the relative strength of your interests in the six different vocational areas.

To discover the types of jobs or careers you might enjoy, look at the first three letters in your code. These letters are numbered "1st", "2nd", and "3rd", and indicate your three strongest areas of interest. The kinds of jobs you might enjoy are likely to be found in those areas.

As you interpret your interest code, keep in mind that the order of the letters in your code is likely to change as your educational background and experiences change. If you are 18 or older, your vocational interests will tend to be relatively stable. If you are 17 or younger, your vocational interests are more likely to change, especially if you have just begun to think about job possibilities. In either case, if you are interested in areas other than the top three areas in your code, you should explore those areas no matter what your code says.

Factors Affecting Your CAT-ASVAB Aptitude Scores

The CAT-ASVAB measures a group of skills that are useful in various training programs and jobs. Your scores on the composite measures reflect the knowledge and skills you have acquired thus far in three general areas. Your scores on the individual tests reflect the knowledge and skills you have acquired thus far in eleven specific areas. Your performance in many of these areas is likely to improve with further training, education, or practical experience.

Your scores depend on many factors including your interests and experiences, and your history of education and training. The scores may not be accurate indicators of your abilities if your schooling, training, or experience has been limited by lack of opportunity, lack of interest, physical handicaps, or other factors. If, for example, you have never worked with automotive and shop equipment, you may be unfamiliar with the terms and concepts covered in the Auto and Shop Information tests. If you were to gain experience or training in those areas, your score on those tests would probably improve.

Gender differences in interests and experiences may also affect performance on some of the CAT-ASVAB tests. Males, for example, often score higher on the Auto, Shop, and Electronics Information tests largely because they tend to acquire more training and experience in those areas. Females, on the other hand, tend to score higher on tests that require rapid decisions and responses, such as the Coding Speed and Numerical Operations tests.

CAT-ASVAB scores may also be affected by factors that prevent a person from answering the questions to the best of his or her ability. Low scores, for example, can often be traced to illness, fatigue, misunderstanding of the directions, distractions during testing, or lack of interest in doing well on the test.

CAT-ASVAB scores may also underestimate your abilities if your native language is not English. Nearly all the tests require use of the English language.

Factors Contributing to Career Success

Aptitude test results are one of many factors you may want to consider when exploring job and career options. Although your scores may suggest promising directions for career exploration, they do not present a complete picture of your chances for success. There are a wide variety of personal attributes—such as motivation, creativity, and determination—that contribute to success in any job or training program. Strengths in these areas can help overcome lack of strong skills in specific areas. You will want to consider these attributes along with other important information—including your interests, past performances, school grades, and personal goals—before making any decisions about your career options.

More About Your Results

More information on how to interpret your results appears on a separate form. Please refer to that form to learn more about the CAT-ASVAB and the Interest Finder, and what they measure.

How to Interpret Your Aptitude Test Results

For the Profile of American Youth Pilot Study, you took twelve tests that make up the Computer Adaptive Test—Armed Services Vocational Aptitude Battery (CAT-ASVAB). The sections below briefly describe the skills the battery is designed to measure. Your scores are reported on a separate form enclosed in this packet.

<p style="text-align: center;">Academic Ability</p> <p>The Academic Ability composite measures how well you did on the Verbal Ability and Math Ability sections combined. It reflects your general potential for formal study and training beyond the high school level.</p>	<p style="text-align: center;">Mathematics Knowledge</p> <p>Mathematics Knowledge scores largely depend on your formal training in mathematics. Most of the questions are about concepts introduced in high school mathematics courses. Low scores in this area may be improved by further training and coursework in mathematics.</p>
<p style="text-align: center;">Verbal Ability</p> <p>Verbal ability measures how well you did on the Word Knowledge and Paragraph Comprehension tests. Your score on this composite is an indicator of your ability to understand and learn from written materials.</p>	<p style="text-align: center;">Electronics Information</p> <p>Electronics Information is a rather specific test. It measures your knowledge of electrical terms and equipment, and your ability to solve electrical problems. Skills in this area are important for people interested in the repair or installation of mechanical equipment. Low scores may be improved through coursework and practical experience.</p>
<p style="text-align: center;">Math Ability</p> <p>Math ability measures how well you did on the Arithmetic Reasoning and Mathematics Knowledge tests. Your score reflects your potential to do well in mathematics courses.</p>	<p style="text-align: center;">Auto and Shop Information</p> <p>The Auto and Shop Information score measures how well you did on the Auto and Shop Information tests combined. It reflects your familiarity with the tools, terms, and practices used in automotive repair, and in metal and wood workshops. Scores on tests of this type are often used to predict success in jobs requiring the repair, maintenance, and operation of mechanical equipment. Low scores may be improved through coursework and hands-on experience.</p>
<p style="text-align: center;">General Science</p> <p>Questions in the General Science test were drawn from a variety of fields including biology, chemistry, and physics. The test measures knowledge of facts and concepts typically taught in high school science courses. Scores on this type of test are often used to predict success in scientific or technical programs. Low scores may be improved through coursework in the sciences, through reading general science publications, or by engaging in scientific hobbies and activities.</p>	<p style="text-align: center;">Mechanical Comprehension</p> <p>Mechanical Comprehension questions included diagrams of mechanical devices such as pulleys, levers, and gears. You were asked questions about how the objects work together. Skills in this area are often important for people interested in mechanical repair, architecture, and engineering. Low scores can probably be improved through practical experience with simple machines.</p>
<p style="text-align: center;">Arithmetic Reasoning</p> <p>Arithmetic Reasoning questions are often called "word problems." They do not require the use of advanced mathematics. Instead, they involve the ability to translate real-life problems into mathematical terms. Skills in this area are important for coursework in subjects such as physics and mathematics. Low scores may be improved by studying applied arithmetic problems.</p>	<p style="text-align: center;">Coding Speed</p> <p>Coding Speed questions asked you to refer to a key to assign numbers to words. The questions required rapid shifts in your attention. They measure the speed and accuracy with which you attend to details. Skills in this area are important for jobs requiring accurate record keeping.</p>
<p style="text-align: center;">Word Knowledge</p> <p>Word Knowledge is a vocabulary test. It measures your understanding of the meanings of words. Your score on this test reflects the variety and quantity of your reading, and your educational experiences. Low scores may be improved by increasing the range and amount of your reading, and by studying vocabulary-building guides.</p>	<p style="text-align: center;">Numerical Operations</p> <p>In the Numerical Operations test you performed arithmetic computations as quickly and accurately as possible. Your performance on the test depends mainly on your speed and accuracy with simple arithmetic operations. Skills in this area are important for jobs requiring accurate record keeping. Low scores may be improved through practice of mental arithmetic.</p>
<p style="text-align: center;">Paragraph Comprehension</p> <p>Paragraph Comprehension measures how well you acquire information from written passages. In this test you were asked to read short paragraphs and answer questions about them. Many jobs and most college courses require this skill to some degree. Low scores in this area may be improved by practice—either by increasing the variety and amount of your reading, or by entering a program especially designed to increase reading speed and comprehension.</p>	<p style="text-align: center;">Assembling Objects</p> <p>The Assembling Objects test was recently added to the battery to measure the ability to mentally rearrange objects in space. The test has not yet been administered to a national sample of individuals. We were unable to compute your percentile score in this area for that reason.</p>

How to Interpret Your Interest Code

The Interest-Finder is designed to identify the types of jobs you might enjoy. You were asked to indicate your likes and dislikes for a wide variety of activities and jobs. Your interest code is based on the preferences you reported. The order of the letters in your code reflects the relative strength of your interests in six vocational areas. Areas toward the beginning of the code are areas in which your interests are strongest. Areas of stronger interest represent promising directions for career exploration.

The sections below briefly describe the six vocational areas in terms of the characteristics and job preferences of people with strong interests in each area. Each area includes many more jobs than are listed below.

Realistic

People with relatively strong interests in the realistic area usually prefer work and activities that are practical and useful. They tend to enjoy mechanical activities and working with their hands. They often prefer to work by themselves rather than with other people. Jobs in the realistic area include:

- aircraft pilot
- butcher
- carpenter
- electrician
- firefighter
- gardener
- inspector
- machinist
- mechanic
- optician
- plumber
- radar operator
- sailor
- truck driver
- woodworker

Social

People whose interests lie in the social area tend to seek jobs in which they can have a positive impact on other people. They typically prefer interacting with people over performing clerical or mechanical tasks. Jobs in this area include:

- counselor
- dental hygienist
- education administrator
- flight attendant
- human resource manager
- occupational therapist
- recreation worker
- religious professional
- social worker
- teacher

Investigative

People in investigative occupations usually like using their knowledge to create new ideas or things. They tend to enjoy scientific and mathematical activities, and learning about new subject matter. Examples of jobs in this area are:

- astronomer
- chemist
- detective
- engineer
- life scientist
- mathematician
- meteorologist
- sociologist
- urban planner

Enterprising

People who prefer enterprising activities typically like public speaking, and managing and leading other people. They usually enjoy being in positions of power and responsibility, and are often found in jobs like the following:

- business executive
- lawyer
- loan officer
- manager
- purchasing agent
- real estate agent
- sales representative
- stock broker
- travel agent

Artistic

Activities that allow creativity are typically preferred by people with strong interests in the artistic area. People with artistic interests usually enjoy using their imagination to create original work. They are often found in the following jobs:

- actor and actress
- architect
- chef
- dancer
- hair stylist
- musician
- news reporter
- photographer
- writer

Conventional

People with strong interests in the conventional area usually prefer activities and jobs that require attention to detail. As a rule, they like performing arithmetic computations and keeping records. Jobs that fall in this area include:

- accountant
- bank teller
- budget analyst
- computer programmer
- court reporter
- librarian
- postal clerk
- secretary
- underwriter

How to Use Your Aptitude Scores and Interest Code

From these descriptions of the aptitude and interest areas, you can see how your results may be helpful in identifying areas of work or study that match your skills and interests. Some of your interests and abilities may lie in different vocational areas. Your score on the Electronics Information test, for example, may indicate that your ability to solve electrical problems is relatively strong, while your interest code may show that you have little interest in jobs that use that ability. Whatever the case may be, you may want to use the results to identify areas where your interests and abilities overlap with the characteristics of jobs or careers.

As you interpret your aptitude and interest scores, keep in mind that they are only one measure of your abilities and interests. Other things you know about yourself—what you like to do, what you do well, your school grades, and other measures—are also indicators of your abilities and interests. When combined with other information, interest and aptitude scores may be useful for helping you decide what area or areas to concentrate on for further study or training, or for job possibilities. Keep in mind that scores on measures of this type are likely to change as your educational background, experiences, and interests change. This is especially true for younger persons in the early stages of their education who have not yet begun to think about career possibilities.

To help you further in thinking about your future, we have enclosed a booklet with questions and answers on careers, jobs, and other tests that you might want to take. Please refer to that booklet for additional information.

Chapter 2 Appendix

Table A.62: NLSY97 Summary Statistics

	Sample Type	Individuals	Mean Age	Male	White	Black	Hispanic
Full Sample	Cross-Section	6,722	14.29	51.28%	69.18%	17.18%	13.64%
	Over-Sample	2,231	14.33	51.01%	0%	56.21%	43.79%
	Total	8984	14.30	51.21%	51.94%	26.91%	21.16%
1997 Sub-sample	Cross-Section	2,598	15.77	49.92%	69.63%	17.01%	13.36%
	Over-Sample	882	15.78	51.93%	0%	57.71%	42.29%
	Total	3,480	15.78	51.98%	27.33%	27.34%	20.69%
2001 Sub-sample	Cross-Section	1,461	18.93	51.68%	67.97%	16.29%	15.74%
	Over-Sample	479	18.99	49.48%	0%	56.16%	43.84%
	Total	1,940	18.94	51.34%	51.19%	26.13%	22.68%
Overlapping	Cross-Section	523	15.81	48.95%	69.02%	15.11%	15.87%
	Over-Sample	187	15.72	49.73%	0%	55.61%	44.39%
	Total	710	15.78	49.15%	50.85%	25.77%	23.38%

Notes: The mean age for the full sample, the 1997 sub-sample, and the overlapping 710 is as of December 31st, 1996. It is the age at interview date for the 2001 sub-sample.

Figure A.4: Distribution of College Completion Expectations in 1997 and 2001

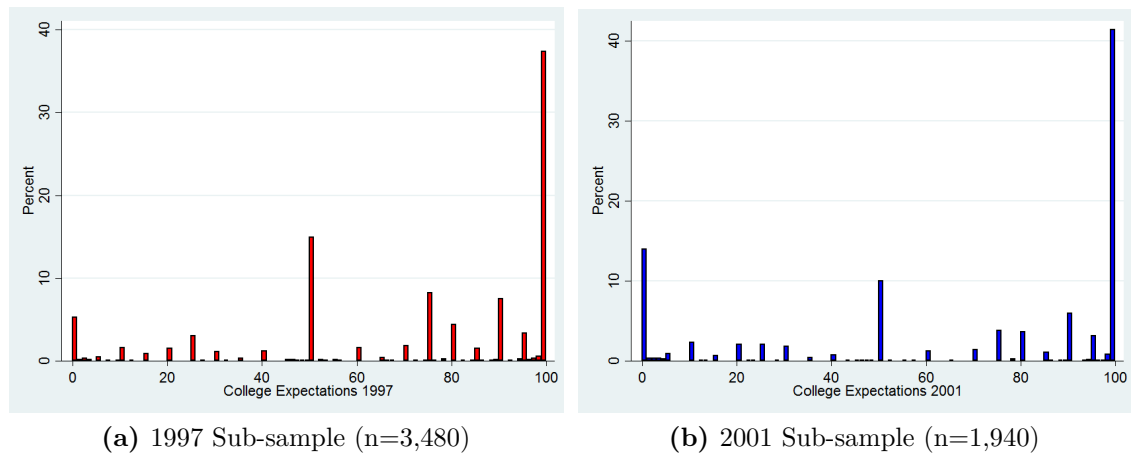


Figure A.5: Distribution of College Completion Expectations Overlapping Sample

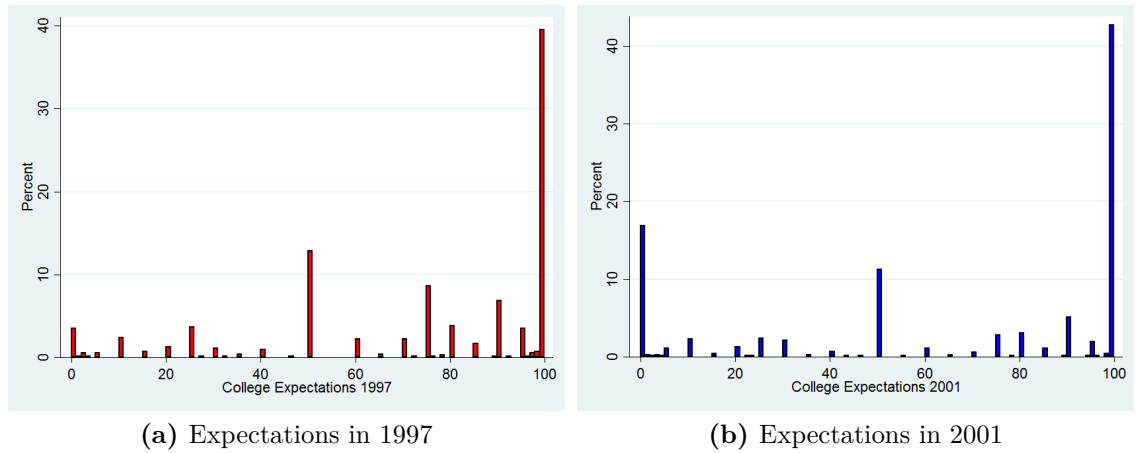


Table A.63: Expectations of Obtaining a Four-year Degree

	N	Mean	Standard Deviation	Minimum	Maximum
1997 Subsample	3,479	72.79%	31.43	0	100
2001 Subsample	1,940	67.17%	38.48	0	100
Overlapping Sample					
In 1997	710	74.29%	30.81	0	100
In 2001	710	64.94%	39.88	0	100
Δ Expectations	710	-9.35	39.49	-100	100

Notes: The standard deviation is reported in percentage points, not percent.

Table A.64: Frequency Distribution of Expectations in 1997 and 2001

	1997 Expectations			2001 Expectations		
	Freq	%	Cum. %	Freq	%	Cum. %
0	25	3.52	3.52	120	16.9	16.9
1	1	0.14	3.66	2	0.28	17.18
2	4	0.56	4.23	1	0.14	17.32
3	1	0.14	4.37	2	0.28	17.61
4	1	0.14	17.75
5	4	0.56	4.93	8	1.13	18.87
10	17	2.39	7.32	16	2.25	21.13
15	5	0.7	8.03	3	0.42	21.55
20	9	1.27	9.3	9	1.27	22.82
22	1	0.14	22.96
23	1	0.14	23.1
25	26	3.66	12.96	17	2.39	25.49
27	1	0.14	13.1
30	8	1.13	14.23	15	2.11	27.61
32	1	0.14	14.37
35	3	0.42	14.79	2	0.28	27.89
40	7	0.99	15.77	5	0.7	28.59
43	1	0.14	28.73
46	1	0.14	15.92	1	0.14	28.87
50	91	12.82	28.73	80	11.27	40.14
55	1	0.14	40.28
60	16	2.25	30.99	8	1.13	41.41
65	3	0.42	31.41	2	0.28	41.69
70	16	2.25	33.66	4	0.56	42.25
72	1	0.14	33.8
75	61	8.59	42.39	20	2.82	45.07
76	1	0.14	42.54
78	2	0.28	42.82	1	0.14	45.21
80	27	3.8	46.62	22	3.1	48.31
85	12	1.69	48.31	8	1.13	49.44
89	1	0.14	48.45	1	0.14	49.58
90	49	6.9	55.35	36	5.07	54.65
92	1	0.14	55.49
94	1	0.14	54.79
95	25	3.52	59.01	14	1.97	56.76
96	1	0.14	59.15	1	0.14	56.9
97	4	0.56	59.72
98	5	0.7	60.42	3	0.42	57.32
99	16	2.25	62.68	19	2.68	60
100	265	37.32	100	284	40	100

Notes: Included above are only the 710 who answered the educational expectations question in both 1997 and 2001. In 1997, students were between 15 and 17 when reporting the percent chance they would complete a four-year college degree. In 2001 they were between 19 and 22.

Figure A.6: Distribution of the Change in College Completion Expectations

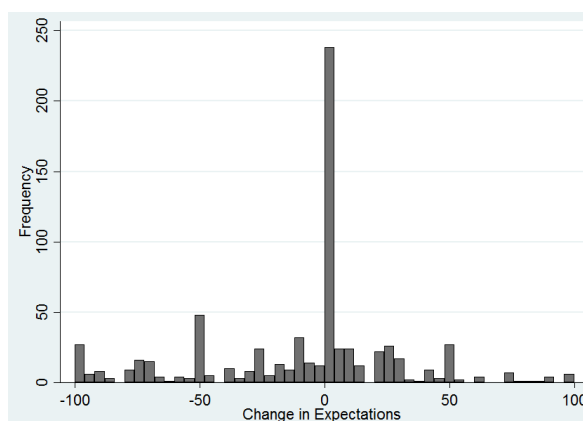


Table A.65: Direction and Magnitude of Updated Expectations from 1997 to 2001

Increase in Chance Have a 4 Year Degree by 30		Number of Individuals	Mean 1997 Expectations	Mean 2001 Expectations	Actual Completion
1 - 25 %		124	76.01%	89.31%	47.58%
26 - 50 %		59	44.49%	86.17%	22.03%
51 - 75 %		13	26.07%	93.08%	7.69%
76 - 90 %		7	8.81%	94.29%	0%
91 - 99 %		1	3%	100%	0%
100 %		5	0%	100%	20%
Total		209	59.56%	89.13%	35.41%
Decrease in Chance Have a 4 Year Degree by 30		Number of Individuals	Mean 1997 Expectations	Mean 2001 Expectations	Actual Completion
1 - 25 %		109	65.02%	51.11%	16.51%
26 - 50 %		74	73.62%	28.08%	0%
51 - 75 %		43	81.93%	12.72%	2.33%
76 - 90 %		20	91.35%	6.75%	0%
91 - 99 %		9	98%	1.89%	0%
100 %		24	100%	0%	4.17%
Total		279	75.87%	29.92%	7.17%
No Change in Chance Have a 4 Year Degree by 30		Number of Individuals	Mean 1997 Expectations	Mean 2001 Expectations	Actual Completion
Total		222	86.17%	86.17%	47.75%

Notes: Included above are only the 710 who answered the educational expectations question in both 1997 and 2001. In 1997, students were between 15 and 17 when reporting the percent chance they will complete a four-year college degree. In 2001 they were between 19 and 22.

Table A.66: Magnitude of Updated Expectations from 1997 to 2001

Increase in Chance				Decrease in Chance			
Have a 4 Year Degree by 30	Number of Individuals	Mean 1997 Expectation	Mean 2001 Expectation	Have a 4 Year Degree by 30	Number of Individuals	Mean 1997 Expectation	Mean 2001 Expectation
1 - 25 %	124	76.01%	89.31%	1 - 25 %	109	65.02%	51.11%
26 - 50 %	59	44.49%	86.17%	26 - 50 %	74	73.62%	28.08%
51 - 75 %	13	26.08%	93.08%	51 - 75 %	43	81.93%	12.72%
76 - 90 %	7	8.14%	94.29%	76 - 90 %	20	91.35%	6.75%
91 - 99 %	1	3%	100%	91 - 99 %	9	98%	1.89%
100 %	5	0%	100%	100 %	24	100%	0%
Total	209	59.56%	89.13%		279	75.87%	29.92%

Notes: 222 individuals did not update their expectations from 1997, for these individuals the Mean is 86.17% with a Standard Deviation equal to 28.92%

Table A.67: Respondents who did not update their expectations from 1997 to 2001

Expectations in 1997 & 2001	Frequency	Percent	Cumulative
0%	14	6.31	6.31%
10%	2	0.9	7.21%
25%	1	0.45	7.66%
30%	1	0.45	8.11%
40%	1	0.45	8.56%
50%	21	9.46	18.02%
75%	5	2.25	20.27%
80%	1	0.45	20.72%
85%	1	0.45	21.17%
90%	7	3.15	24.32%
95%	1	0.45	24.77%
99%	1	0.45	25.23%
100%	166	74.77	100%
Total	222	100	

Notes:

Table A.68: Uncertainty Estimates

	<u>Probit</u>	<u>Y= Δ Expectations_i </u>	
	(1)	(2)	(3)
<i>Expect_i</i> ∈ [25, 75]	0.921*** (0.120)	8.491*** (2.106)	5.071 (2.656)
Observations	710	710	420
Adjusted <i>R</i> ²		0.016	0.007

Notes: Column (3) contains estimates using those who reported between a 1% and 99% chance initially.

– * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

– Robust standard errors in parenthesis

Table A.69: Uncertainty Estimates

	Probit	Y= $ \Delta \text{ Expectations}_i $					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\text{Expect}_i \in [25, 75]$	0.921*** (0.120)	8.491*** (2.106)	5.071 (2.657)	5.071 (2.656)			
$(50 - \text{Expectation}_{1997,i})^2$					-0.00460*** (0.00106)	-0.00351 (0.00185)	-0.00351 (0.00185)
$\text{Expectations}_{1997} = 0$			2.432 (8.214)			5.301 (8.708)	
$\text{Expectations}_{1997} = 100$			-6.360* (3.097)			-3.491 (4.238)	
Observations	710	710	710	420	710	710	420
Adjusted R^2		0.016	0.021	0.007	0.021	0.022	0.008

Notes: Columns (3) and (7) contain estimates using only those who reported between a 1% and 99% chance initially.

- * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

- Robust standard errors in parenthesis

Table A.70: Summary statistics for the Indicators and Information Sources

Variable	Mean	Std. Dev.
Expectations in 1997	74.289	30.807
Expectations in 2001	64.937	39.883
Change in Expectations	-9.352	39.494
$\text{Expectations}_{1997} = 0$	0.035	0.184
$\text{Expectations}_{1997} = 100$	0.373	0.484
$\text{Expectations}_{1997} \in [25, 75]$	0.331	0.471
ASVAB Percentile	45.862	26.948
Took the ACT	0.307	0.462
Took the SAT	0.187	0.39
Took the ACT or SAT	0.434	0.496
Enrolled in a 2yr College in 2001	0.123	0.328
Enrolled in a 4yr College in 2001	0.273	0.446
N	710	

Notes: The ASVAB Percentile is a semi-continuous measure that ranges from 0 to 100. The ACT/SAT and Enrollment variables are indicators of whether or not they took an exam or were enrolled in either institution when they answered the expectations questions in 2001.

Table A.71: Summary Statistics of Information Sources

Variable	Mean	Std. Dev.
ASVAB Percentile	45.862	26.948
Took the ACT	0.307	0.462
Took the SAT	0.187	0.39
Took the ACT or SAT	0.434	0.496
Enrolled in a 2yr College in 2001	0.123	0.328
Enrolled in a 4yr College in 2001	0.273	0.446
N	710	

Notes: The ASVAB Percentile is a semi-continuous measure that ranges from 0 to 100. The ACT/SAT and Enrollment variables are indicators of whether or not they took an exam or were enrolled in either institution when they answered the expectations questions in 2001.

Table A.72: 2001 Expectation Formation - Interior Reporters

Y = Expectation of completing a college degree in 2001						
	(1)	(2)	(3)	(4)	(5)	(6)
College Expectations 1997	0.540*** (0.0588)	0.794*** (0.0377)	0.668*** (0.0358)	0.849*** (0.0282)	0.523*** (0.0611)	0.478*** (0.0607)
ASVAB Percentile	0.507*** (0.0723)				0.487*** (0.0756)	0.335*** (0.0757)
Took the ACT or SAT Ever		11.95** (4.160)			3.129 (4.625)	-1.356 (4.344)
Enrolled in 2yr College			33.47*** (5.702)			31.05*** (6.142)
Enrolled in 4yr College			41.53*** (3.332)			33.17*** (4.271)
Change As				-1.917 (13.13)	-9.158 (12.73)	-9.602 (10.37)
Change As & Bs				-14.38 (12.78)	-19.38 (12.36)	-16.00 (10.31)
Change Bs				-17.20 (13.12)	-20.78 (12.55)	-19.05 (10.35)
Change Bs & Cs				-9.285 (12.92)	-18.06 (12.16)	-16.49 (10.04)
Change Cs				-6.669 (13.22)	-13.32 (12.70)	-11.01 (10.91)
Change Cs & Ds				-13.30 (13.65)	-21.16 (12.84)	-18.95 (11.12)
Change Ds				-5.271 (15.99)	-9.445 (15.66)	-6.983 (13.53)
Change Ds & other				0.246 (18.76)	-5.167 (17.99)	-0.904 (16.19)
Observations	341	420	420	420	341	341
Adjusted R^2	0.748	0.701	0.750	0.697	0.749	0.783

Notes: Only those who reported between 1% and 99% in the initial round were used to obtain these estimates. When the ASVAB percentile is used only those without imputed ASVAB scores were utilized.

- Robust standard errors in parenthesis

-* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.73: Grade Change Matrix

			<u>High School Grades</u>							
			As n=70	As and Bs n=132	Bs n=133	Bs and Cs n=186	Cs n=95	Cs and Ds n=52	Ds n=12	Fs n=8
8th Grade <u>Grades</u>	As	n=107	51	34	18	3	1	0	0	0
	As and Bs	n=152	14	47	45	34	7	4	0	1
	Bs	n=94	1	22	33	29	7	2	0	0
	Bs and Cs	n=162	4	18	23	64	31	18	0	4
	Cs	n=100	0	8	12	32	30	13	5	0
	Cs and Ds	n=48	0	2	2	20	13	8	1	2
	Ds	n=18	0	1	0	4	6	4	3	0
	Fs	n=7	0	0	0	0	0	3	3	1

Notes: 17 individuals that reported "other grades" in 8th grade and/or high school are not included in this table.

Table A.74: Grade Change Matrix

			<u>High School Grades</u>							
			As	As and Bs	Bs	Bs and Cs	Cs	Cs and Ds	Ds	Less Ds
8th Grade	As	0	6.39	9.45	6.89	1.41	9.34	-2.62	-8.70	
		51	34	18	3	1	0	0	0	
	As and Bs	-6.39	0	3.05	0.50	-4.99	2.95	-9.01	-15.09	
		14	47	45	34	7	4	0	1	
	Bs	-9.45	-3.05	0	-2.55	-8.04	-0.10	-12.07	-18.14	
		1	22	33	29	7	2	0	0	
	Bs and Cs	-6.89	-0.50	2.55	0	-5.48	2.45	-9.51	-15.59	
		4	18	23	64	31	18	0	4	
	Cs	-1.41	4.99	8.04	5.48	0	7.94	-4.03	-10.11	
		0	8	12	32	30	13	5	0	
	Cs and Ds	-9.34	-2.95	0.10	-2.45	-7.94	0	-11.96	-18.04	
		0	2	2	20	13	8	1	2	
	Ds	2.62	9.01	12.07	9.51	4.03	11.96	0	-6.08	
		0	1	0	4	6	4	3	0	
	Less Ds	8.70	15.09	18.14	15.59	10.11	18.04	6.08	0	
		0	0	0	0	0	3	3	1	
			Grades Decreased from 8th grade to High School							
			Grades Increased from 8th grade to High School							
			Grades were the same in 8th grade and High school							

Notes: The predicted change in expectations for each grade change combination is bold and underneath each are the number of respondents who are observed making those changes.

-This includes estimates from only those who reported between 1% and 99% initially.

-17 individuals that reported "other grades" in 8th grade and/or high school are not included.

Table A.75: Model with Variable Descriptions

<u>Variable of Interest:</u>	The percent chance the student believes that they will obtain a 4 year college degree by the time they are 30	
<u>Control Variables:</u>	Demographics - Male - White - Black Family - Only child - Household members under 18 - Mother, Father, or both parents absent - Income quantile Location - Urban - MSA Central City - MSA Non Central City - 3 Census Tracts (W, NE, NC)	Parent Education - One parent has some college - Both parents have some college - One parent has a college degree - Both parents have a college degree Schooling: - Self-reported grades in 8th grade (levels) - Self-reported grades in HS if applicable (levels) -ASVAB Percentile -Enrollment in high school indicator Other: - % of peers that plan to go to college - Teacher quality and involvement indicators

Notes: For each categorical variable, that which is not listed is used as the base case. For grade controls, if a student is still in high school then the 8th grade controls are used, when they have graduated or left school the high school grade controls are used. In each case the student reported their grades in levels. Transcript data for high school GPA is available but low reporting severely reduces sample size. Estimates from specifications that have included it do not differ drastically. The teacher quality and involvement controls are reported in levels by the student. The student reports whether they believe their teachers overall were 'good' and whether they were 'interested in the students success.' They responded with strongly agree, agree, disagree, or strongly disagree.

Table A.76: 1997 Expectation Formation

<u>Y = Expectation of completing a college degree in 1997</u>						
	(1)	(2)	(3)	(4)	(5)	(6)
Parent's College Expectations				0.416*** (0.0207)		0.420*** (0.0479)
Age in 1997	-2.952*** (0.782)	-3.097*** (0.767)	-3.573*** (0.845)	-1.265 (0.834)	-0.373 (1.903)	0.931 (1.861)
Male	-8.144*** (1.050)	-8.726*** (1.029)	-4.645*** (0.963)	-4.944*** (0.946)	-4.356* (2.163)	-3.801 (2.127)
White	5.691*** (1.616)	0.259 (1.732)	-5.266** (1.601)	-1.544 (1.555)	-6.325 (3.622)	-5.437 (3.525)
Black	6.884*** (1.551)	7.652*** (1.684)	6.617*** (1.550)	4.458** (1.570)	6.176 (3.374)	2.119 (3.408)
Income Quantile		6.361*** (0.637)	0.914 (0.628)	-0.169 (0.589)	1.087 (1.395)	-0.00571 (1.249)
As in 8th			17.65*** (1.858)	10.07*** (1.888)	20.83*** (3.940)	12.67*** (3.719)
As & Bs in 8th			13.04*** (1.801)	7.263*** (1.813)	19.20*** (3.933)	11.84** (3.780)
Bs in 8th			10.80*** (1.928)	7.653*** (1.944)	15.92*** (4.314)	9.012* (4.319)
Bs & Cs in 8th			3.573* (1.802)	2.434 (1.788)	7.683 (3.990)	6.046 (3.754)
Cs & Ds in 8th			-5.696* (2.446)	-1.445 (2.463)	-3.289 (5.635)	-0.253 (5.205)
Ds in 8th			-5.797 (3.607)	0.783 (3.481)	0.277 (6.228)	3.017 (5.888)
Fs in 8th			-8.323 (5.815)	-1.683 (5.093)	-28.43** (10.15)	-12.45 (12.48)
No HS Degree for both Parents			-3.544 (2.024)	-4.247* (2.066)	-1.458 (4.677)	-2.454 (4.439)
One Parent has Some College			5.111*** (1.170)	1.961 (1.158)	5.276* (2.612)	4.315 (2.680)
Both Parents have Some College			3.774* (1.915)	4.580** (1.775)	5.727 (3.981)	0.458 (4.095)
One Parent has a College Degree			10.64*** (1.246)	7.322*** (1.187)	10.89*** (2.949)	6.819* (2.825)
Both Parents have a College Degree			2.715 (1.480)	0.950 (1.389)	4.267 (3.226)	3.180 (3.096)
ASVAB Percentile			0.119*** (0.0227)	0.0346 (0.0224)	0.102 (0.0531)	0.0174 (0.0545)
Highest Grade Completed 1997			2.376*** (0.650)	0.843 (0.651)	0.721 (1.504)	-0.983 (1.443)
Enrolled in HS in 1997			24.16*** (2.678)	14.23*** (2.655)	25.16*** (6.787)	9.801 (5.990)
Constant	116.3*** (12.38)	107.7*** (12.50)	65.95*** (12.28)	33.64** (11.82)	23.38 (28.58)	16.97 (28.30)
Location Controls	No	Yes	Yes	Yes	Yes	Yes
Family Structure Controls	No	Yes	Yes	Yes	Yes	Yes
Observations	3479	3389	3386	2905	689	601
Adjusted R^2	0.030	0.086	0.255	0.389	0.248	0.357

Notes: Location controls include: whether or not the family lived in an MSA and if so if it were in the central city or not; whether they lived in an urban or rural area; and which of the four large standard census regions they lived. Controls for household structure include: whether the mother, father, or both were absent from the household; if the student were an only child; and the number of other siblings under 18 in the household in 1997.

—* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

—Robust standard errors in parentheses

Table A.77: 1997 Expectation Formation and Completion

	Y = Expectation of completing a college degree in 1997				Probit for Completion of College Degree	
	(1)	(2)	(3)	(4)	(5)	(6)
Parent's College Expectations		0.416*** (0.0207)		0.420*** (0.0479)	0.0114*** (0.00135)	0.0152*** (0.00318)
Age in 1997	-3.573*** (0.845)	-1.265 (0.834)	-0.373 (1.903)	0.931 (1.861)	-0.0799 (0.0541)	0.115 (0.121)
Male	-4.645*** (0.963)	-4.944*** (0.946)	-4.356* (2.163)	-3.801 (2.127)	-0.118 (0.0617)	-0.150 (0.144)
White	-5.266** (1.601)	-1.544 (1.555)	-6.325 (3.622)	-5.437 (3.525)	0.0210 (0.110)	0.143 (0.233)
Black	6.617*** (1.550)	4.458** (1.570)	6.176 (3.374)	2.119 (3.408)	0.337** (0.110)	0.442 (0.249)
Income Quintile	0.914 (0.628)	-0.169 (0.589)	1.087 (1.395)	-0.00571 (1.249)	0.114** (0.0395)	0.0295 (0.0901)
As in 8th	17.65*** (1.858)	10.07*** (1.888)	20.83*** (3.940)	12.67*** (3.719)	0.923*** (0.130)	1.254*** (0.292)
As & Bs in 8th	13.04*** (1.801)	7.263*** (1.813)	19.20*** (3.933)	11.84** (3.780)	0.490*** (0.118)	0.490 (0.264)
Bs in 8th	10.80*** (1.928)	7.653*** (1.944)	15.92*** (4.314)	9.012* (4.319)	0.358** (0.125)	0.534 (0.282)
Bs & Cs in 8th	3.573* (1.802)	2.434 (1.788)	7.683 (3.990)	6.046 (3.754)	0.0875 (0.121)	0.0610 (0.264)
Cs & Ds in 8th	-5.696* (2.446)	-1.445 (2.463)	-3.289 (5.635)	-0.253 (5.205)	-0.172 (0.195)	-0.244 (0.438)
No HS Degree for both Parents	-3.544 (2.024)	-4.247* (2.066)	-1.458 (4.677)	-2.454 (4.439)	-0.0393 (0.144)	0.154 (0.308)
One Parent has Some College	5.111*** (1.170)	1.961 (1.158)	5.276* (2.612)	4.315 (2.680)	0.0375 (0.0722)	0.182 (0.162)
Both Parents have Some College	3.774* (1.915)	4.580** (1.775)	5.727 (3.981)	0.458 (4.095)	0.233 (0.124)	0.0674 (0.282)
One Parent has a College Degree	10.64*** (1.246)	7.322*** (1.187)	10.89*** (2.949)	6.819* (2.825)	0.465*** (0.0799)	0.827*** (0.193)
Both Parents have a College Degree	2.715 (1.480)	0.950 (1.389)	4.267 (3.226)	3.180 (3.096)	0.257* (0.117)	-0.114 (0.276)
ASVAB Percentile	0.119*** (0.0227)	0.0346 (0.0224)	0.102 (0.0531)	0.0174 (0.0545)	0.0104*** (0.00143)	0.00779* (0.00338)
Highest Grade Completed 1997	2.376*** (0.650)	0.843 (0.651)	0.721 (1.504)	-0.983 (1.443)	0.100* (0.0436)	0.0969 (0.0962)
Enrolled in HS in 1997	24.16*** (2.678)	14.23*** (2.655)	25.16*** (6.787)	9.801 (5.990)	0.257 (0.269)	0.514 (0.487)
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes
Family Structure Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3386	2905	689	601	2905	601
Adjusted R^2	0.255	0.389	0.248	0.357		

Notes: Location controls include: whether or not the family lived in an MSA and if so if it were in the central city or not; whether they lived in an urban or rural area; and which of the four large standard census regions they lived. Controls for household structure include: whether the mother, father, or both were absent from the household; if the student were an only child; and the number of other siblings under 18 in the household in 1997.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

—Robust standard errors in parentheses

Table A.78: 2001 Expectation Formation - Full Controls

	Expectations in 1997	Y = Expectation of completing a college degree in 2001				
		(2)	(3)	(4)	(5)	(6)
College Expectations 1997					0.125* (0.0529)	0.169** (0.0622)
Parent's College Expectations						0.0213 (0.0599)
Age in 1997	-3.573*** (0.845)	-4.320*** (1.157)	-2.406** (0.777)	-2.407 (1.788)	-2.175 (1.782)	-3.181 (1.939)
Male	-4.645*** (0.963)	-4.775** (1.584)	-2.701 (1.473)	-2.594 (2.547)	-2.127 (2.534)	-2.631 (2.712)
White	-5.266** (1.601)	-5.570* (2.673)	-3.596 (2.459)	-2.292 (4.416)	-1.781 (4.385)	-1.442 (4.711)
Black	6.617*** (1.550)	5.229* (2.626)	3.657 (2.442)	4.607 (4.129)	3.859 (4.152)	2.336 (4.543)
Income Quantile	0.914 (0.628)	2.423* (1.040)	1.094 (0.975)	1.351 (1.730)	1.351 (1.709)	1.127 (1.780)
No HS Degree for both Parents	-3.544 (2.024)	-4.372 (3.155)	-3.051 (2.907)	-1.687 (4.950)	-1.386 (4.979)	0.0640 (5.494)
One Parent has Some College	5.111*** (1.170)	4.592* (2.068)	2.155 (1.893)	-1.707 (3.107)	-2.173 (3.070)	-1.440 (3.310)
Both Parents have Some College	3.774* (1.915)	10.27** (3.199)	10.52*** (2.981)	11.37* (5.350)	10.56* (5.294)	9.897 (5.373)
One Parent has a College Degree	10.64*** (1.246)	13.92*** (2.194)	10.56*** (2.061)	10.70** (3.554)	9.708** (3.512)	8.684* (3.624)
Both Parents have a College Degree	2.715 (1.480)	4.408 (2.606)	1.243 (2.459)	-2.618 (4.341)	-3.031 (4.269)	-1.552 (4.480)
ASVAB Percentile	0.119*** (0.0227)	0.246*** (0.0409)	0.141*** (0.0389)	0.180** (0.0676)	0.176** (0.0675)	0.169* (0.0753)
Highest Grade Completed 1997	2.376*** (0.650)	2.896** (1.106)				
Enrolled in HS in 1997	24.16*** (2.678)	11.16 (6.581)				
As in HS			7.194* (3.574)	-4.727 (5.898)	-4.377 (5.706)	-2.416 (6.127)
As & Bs in HS			8.475** (3.104)	6.646 (5.332)	7.106 (5.201)	7.597 (5.537)
Bs in HS			4.772 (2.991)	1.795 (4.941)	2.535 (4.840)	3.962 (5.086)
Bs & Cs in HS			3.741 (2.778)	4.851 (4.384)	5.578 (4.286)	7.366 (4.548)
Cs & Ds in HS			-0.733 (3.755)	-0.544 (5.957)	-0.518 (6.040)	-3.900 (6.505)
Ds in HS			-3.762 (6.886)	-0.463 (12.48)	-0.288 (13.37)	3.304 (17.88)
Fs in HS			-3.457 (5.853)	-0.426 (8.919)	1.674 (8.629)	5.932 (9.893)
Other Grades in 8th			-2.092 (5.160)	-4.554 (9.431)	-4.783 (9.375)	1.430 (8.989)
Highest Grade Completed 2001			4.964*** (0.837)	4.635*** (1.368)	4.177** (1.360)	3.782* (1.475)
Enrolled in HS 2001			18.83*** (2.470)	9.841 (14.50)	11.16 (15.38)	14.60 (16.14)
Enrolled in 2yr College			19.47*** (3.141)	24.37*** (4.414)	23.56*** (4.388)	21.89*** (4.906)
Enrolled in 4yr College			27.02*** (2.203)	34.43*** (3.553)	33.44*** (3.510)	33.72*** (3.780)
Constant	65.95*** (12.28)	70.25*** (14.24)	11.03 (12.44)	9.491 (30.31)	4.346 (30.22)	22.04 (33.77)
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes
Family Structure Controls	Yes	Yes	Yes	Yes	Yes	Yes
Grades in 8th Grade	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3386	1893	1894	690	690	602
Adjusted R ²	0.255	0.248	0.357	0.400	0.406	0.404

Notes: This model includes all of the new information available to students but does not look at how things change. Location controls are for as of 2001 and include: whether or not the family lived in an MSA and if so if it were in the central city or not; whether they lived in an urban or rural area; and which of the four large standard census regions they lived. Controls for household structure include: whether the mother, father, or both were absent from the household; if the student were an only child; and the number of other siblings under 18 in the household in 1997.

-* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

-Robust standard errors in parentheses

Table A.79: Change in Expectations: Overall and Magnitude

	$Y = \Delta$ in Expectations			$Y = \Delta $ in Expectations		
	(1)	(2)	(3)	(4)	(5)	(6)
College Expectations 1997		-0.875*** (0.0529)	-0.831*** (0.0622)		0.0721 (0.0551)	0.0725 (0.0625)
Parent's College Expectations			0.0213 (0.0599)			-0.00797 (0.0524)
Age in 1997	-0.545 (2.269)	-2.175 (1.782)	-3.181 (1.939)	1.498 (1.676)	1.633 (1.680)	2.737 (1.813)
Male	1.148 (3.139)	-2.127 (2.534)	-2.631 (2.712)	2.556 (2.249)	2.826 (2.234)	4.134 (2.383)
White	1.805 (5.291)	-1.781 (4.385)	-1.442 (4.711)	8.757* (3.790)	9.052* (3.766)	10.79** (3.974)
Black	-1.393 (5.090)	3.859 (4.152)	2.336 (4.543)	4.292 (3.544)	3.860 (3.584)	5.717 (3.893)
Income Quantile	1.350 (2.029)	1.351 (1.709)	1.127 (1.780)	1.624 (1.471)	1.624 (1.464)	1.665 (1.540)
No HS Degree for both Parents	0.733 (6.628)	-1.386 (4.979)	0.0640 (5.494)	3.579 (4.449)	3.753 (4.384)	2.298 (4.634)
One Parent has Some College	-5.442 (3.653)	-2.173 (3.070)	-1.440 (3.310)	-4.425 (2.657)	-4.694 (2.673)	-3.277 (2.888)
Both Parents have Some College	4.867 (6.155)	10.56* (5.294)	9.897 (5.373)	-1.785 (4.451)	-2.254 (4.589)	-3.901 (4.858)
One Parent has a College Degree	2.708 (4.073)	9.708** (3.512)	8.684* (3.624)	-7.235* (2.894)	-7.812* (3.046)	-8.054** (3.099)
Both Parents have a College Degree	-5.935 (4.788)	-3.031 (4.269)	-1.552 (4.480)	-4.262 (3.839)	-4.501 (3.882)	-2.848 (4.099)
Enrolled in HS 2001	20.38 (23.86)	11.16 (15.38)	14.60 (16.14)	15.18 (13.62)	15.94 (13.67)	0.618 (15.17)
Enrolled in 2yr College	17.85*** (5.221)	23.56*** (4.388)	21.89*** (4.906)	-11.38** (3.849)	-11.85** (3.992)	-12.62** (4.310)
Enrolled in 4yr College	26.50*** (3.938)	33.44*** (3.510)	33.72*** (3.780)	-23.75*** (3.079)	-24.32*** (3.270)	-24.09*** (3.524)
ASVAB Percentile	0.151 (0.0835)	0.176** (0.0675)	0.169* (0.0753)	-0.102 (0.0597)	-0.104 (0.0597)	-0.105 (0.0660)
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes
School Type and Completion Controls	Yes	Yes	Yes	Yes	Yes	Yes
Grades in 8th Grade and HS	Yes	Yes	Yes	Yes	Yes	Yes
Family Structure Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	690	690	602	690	690	602
Adjusted R^2	0.071	0.404	0.378	0.206	0.208	0.215

Notes: OLS estimates presented. See the variable description table above for a list of the controls for location, school type, grades, and family structure.

-* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

-Robust standard errors in parentheses

Table A.80: Change in Expectations: Any, Non-positive, and Negative

	Y=1 if $\Delta \text{Exp} \neq 0$			Y=1 if $\Delta \text{Exp} \geq 0$			Y=1 if $\Delta \text{Exp} > 0$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
College Expectations 1997		-0.00246** (0.000832)	-0.00286** (0.000952)		-0.00642*** (0.000832)	-0.00647*** (0.000976)		-0.00830*** (0.000772)	-0.00859*** (0.000920)
Parent's College Expectations			0.000658 (0.000905)			0.00126 (0.000906)			0.00172 (0.000885)
Age in 1997	0.0220 (0.0280)	0.0171 (0.0284)	0.0268 (0.0306)	-0.00422 (0.0294)	-0.0196 (0.0307)	-0.0227 (0.0338)	0.0105 (0.0265)	-0.00338 (0.0264)	-0.00289 (0.0278)
Male (d)	0.118** (0.0382)	0.112** (0.0386)	0.125** (0.0419)	-0.0478 (0.0428)	-0.0601 (0.0443)	-0.0602 (0.0480)	0.0728 (0.0374)	0.0498 (0.0382)	0.0594 (0.0404)
White (d)	0.0460 (0.0637)	0.0331 (0.0643)	0.000259 (0.0684)	0.0242 (0.0652)	0.000211 (0.0673)	0.0294 (0.0720)	0.0657 (0.0591)	0.0369 (0.0604)	0.0336 (0.0628)
Black (d)	-0.112 (0.0653)	-0.100 (0.0649)	-0.128 (0.0709)	-0.00768 (0.0652)	0.0398 (0.0651)	0.0556 (0.0703)	-0.102 (0.0543)	-0.0643 (0.0552)	-0.0845 (0.0564)
Income Quantile	0.0281 (0.0251)	0.0297 (0.0253)	0.0262 (0.0268)	0.0158 (0.0282)	0.0244 (0.0291)	0.0195 (0.0308)	0.0421 (0.0247)	0.0495 (0.0259)	0.0433 (0.0273)
Siblings under 18	0.0292 (0.0157)	0.0320* (0.0155)	0.0351* (0.0169)	0.0129 (0.0163)	0.0188 (0.0168)	0.0179 (0.0184)	0.0361* (0.0145)	0.0506*** (0.0144)	0.0560*** (0.0157)
Mother Absent (d)	0.186* (0.0745)	0.192** (0.0701)	0.206** (0.0653)	-0.142 (0.139)	-0.109 (0.136)	-0.120 (0.146)	0.0877 (0.121)	0.110 (0.140)	0.120 (0.152)
Father Absent (d)	0.127** (0.0479)	0.119* (0.0485)	0.124* (0.0511)	0.00739 (0.0565)	-0.00769 (0.0593)	-0.00671 (0.0628)	0.146** (0.0544)	0.131* (0.0571)	0.139* (0.0594)
Both Parents Absent (d)	0.0409 (0.0903)	0.0638 (0.0862)	0.00782 (0.0996)	-0.253* (0.118)	-0.245 (0.132)	-0.195 (0.142)	-0.182** (0.0692)	-0.194*** (0.0506)	-0.208*** (0.0430)
No HS Degree for both Parents (d)	-0.0229 (0.0725)	-0.0331 (0.0751)	-0.0651 (0.0825)	0.0281 (0.0701)	0.0168 (0.0746)	0.0768 (0.0765)	-0.00977 (0.0647)	-0.0346 (0.0650)	-0.00678 (0.0730)
One Parent has Some College (d)	-0.118* (0.0495)	-0.108* (0.0493)	-0.0847 (0.0524)	-0.0178 (0.0518)	-0.00126 (0.0531)	-0.000365 (0.0565)	-0.124** (0.0426)	-0.106* (0.0426)	-0.0863 (0.0447)
Both Parents have Some College (d)	0.0492 (0.0689)	0.0653 (0.0651)	0.0455 (0.0716)	0.0446 (0.0840)	0.104 (0.0810)	0.0963 (0.0869)	0.0620 (0.0881)	0.147 (0.0999)	0.137 (0.105)
One Parent has a College Degree (d)	-0.0408 (0.0577)	-0.0253 (0.0564)	-0.0278 (0.0616)	0.0450 (0.0624)	0.109 (0.0601)	0.0815 (0.0662)	-0.0145 (0.0550)	0.0663 (0.0595)	0.0376 (0.0584)
Both Parents have a College Degree (d)	-0.153 (0.0934)	-0.143 (0.0916)	-0.150 (0.0991)	-0.0420 (0.102)	-0.0372 (0.107)	0.0268 (0.107)	-0.127* (0.0626)	-0.115* (0.0576)	-0.0892 (0.0619)
Enrolled in HS 2001 (d)	0.0416 (0.205)	0.0307 (0.220)	-0.116 (0.269)	0.0919 (0.158)	0.0175 (0.165)	0.0836 (0.176)	0.194 (0.189)	0.0746 (0.172)	0.0355 (0.190)
Enrolled in 2yr College (d)	-0.167* (0.0691)	-0.145* (0.0671)	-0.141 (0.0739)	0.171** (0.0524)	0.211*** (0.0494)	0.200*** (0.0553)	0.0589 (0.0657)	0.126 (0.0700)	0.102 (0.0752)
Enrolled in 4yr College (d)	-0.337*** (0.0648)	-0.319*** (0.0637)	-0.327*** (0.0685)	0.459*** (0.0388)	0.492*** (0.0370)	0.490*** (0.0410)	0.127* (0.0594)	0.215*** (0.0618)	0.184** (0.0640)
ASVAB Percentile	0.00120 (0.00100)	0.00135 (0.00101)	0.00107 (0.00107)	0.000960 (0.00106)	0.00104 (0.00108)	0.00127 (0.00117)	0.00240* (0.000968)	0.00299** (0.000965)	0.00280** (0.00102)
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Type and Completion Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grades in 8th Grade and HS	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	678	678	591	686	686	599	690	690	599
Pseudo R^2	0.175	0.186	0.191	0.203	0.273	0.283	0.093	0.252	0.261

Notes: OLS was used in the estimation above. Marginal Effects from a probit model were similar.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

-Robust standard errors in parentheses

Table A.81: Summary Statistics of Bayesian Components

	Variable	Mean	Std. Dev.	N
Full Sample	$Prob(S_{c,2001} ASVAB_i)$	0.649	0.399	710
	$Prob(S_{c,1997})$	0.743	0.308	710
	$Prob(ASVAB_i S_c)$	0.264	0.204	710
	$Prob(ASVAB_i)$	0.423	0.222	710
	$\frac{Prob(ASVAB_i S_c)}{Prob(ASVAB_i)}$	0.607	0.405	706
Interior Sample	$Prob(S_{c,2001} ASVAB_i)$	0.582	0.399	420
	$Prob(S_{c,1997})$	0.625	0.268	420
	$Prob(ASVAB_i S_c)$	0.253	0.2	420
	$Prob(ASVAB_i)$	0.479	0.174	420
	$\frac{Prob(ASVAB_i S_c)}{Prob(ASVAB_i)}$	0.399	0.339	420

Notes: The interior sample are those who reported between 1% and 99% in 1997.

Table A.82: Bayesian Component Estimation

ASVAB Percentile	Number of Students	Number that Complete College	\$Prob(ASVAB)\$	\$Prob(ASVAB S_c)\$	Information Processing
0	62	0	0.006925827	0	0
1	95	1	0.010612154	0.000438789	0.041347775
2	104	0	0.011617516	0	0
3	88	2	0.009830206	0.000877578	0.089273605
4	90	3	0.010053619	0.001316367	0.13093462
5	86	2	0.009606792	0.000877578	0.091349735
6	97	1	0.010835567	0.000438789	0.040495243
7	90	5	0.010053619	0.002193945	0.218224367
8	91	3	0.010165326	0.001316367	0.129495778
9	87	1	0.009718499	0.000438789	0.045149869
10	91	7	0.010165326	0.003071523	0.302156816
11	102	3	0.011394102	0.001316367	0.115530547
12	89	3	0.009941912	0.001316367	0.132405796
13	90	7	0.010053619	0.003071523	0.305514114
14	95	7	0.010612154	0.003071523	0.289434424
15	92	6	0.010277033	0.002632734	0.256176431
16	79	3	0.008824844	0.001316367	0.149166023
17	86	6	0.009606792	0.002632734	0.274049206
18	85	7	0.009495085	0.003071523	0.323485533
19	89	10	0.009941912	0.004387889	0.441352653
20	82	10	0.009159964	0.004387889	0.479029099
21	83	13	0.009271671	0.005704256	0.615234964
22	80	10	0.00893655	0.004387889	0.491004827
23	85	14	0.009495085	0.006143045	0.646971066
24	82	12	0.009159964	0.005265467	0.574834919
25	80	12	0.00893655	0.005265467	0.589205792
26	77	9	0.00860143	0.0039491	0.459121396
27	366	39	0.040884718	0.017112769	0.418561492
28	80	16	0.00893655	0.007020623	0.785607723
29	73	15	0.008154602	0.006581834	0.807131222
30	79	13	0.008824844	0.005704256	0.646386101
31	81	9	0.009048257	0.0039491	0.436448735
32	317	54	0.035411081	0.023694603	0.669129606
33	66	11	0.007372654	0.004826678	0.654673102
34	71	7	0.007931189	0.003071523	0.387271413
35	353	42	0.039432529	0.018429136	0.467358702
36	325	49	0.036304736	0.021500658	0.59222736
37	74	12	0.008266309	0.005265467	0.636979235
38	69	14	0.007707775	0.006143045	0.796993342
39	62	10	0.006925827	0.004387889	0.633554615
40	73	21	0.008154602	0.009214568	1.129983711
41	64	10	0.00714924	0.004387889	0.613756033
42	78	12	0.008713137	0.005265467	0.604313633
43	70	15	0.007819482	0.006581834	0.84172256
44	69	15	0.007707775	0.006581834	0.853921438
45	72	9	0.008042895	0.0039491	0.491004827
46	69	25	0.007707775	0.010969724	1.423202396
47	72	22	0.008042895	0.009653357	1.200234021
48	68	23	0.007596068	0.010092146	1.328601296
49	67	18	0.007484361	0.007898201	1.055293956
50	63	19	0.007037534	0.00833699	1.184646566
51	71	18	0.007931189	0.007898201	0.995840775
52	63	19	0.007037534	0.00833699	1.184646566
53	64	22	0.00714924	0.009653357	1.350263273
54	63	17	0.007037534	0.007459412	1.059946927
55	60	15	0.006702413	0.006581834	0.982009653
56	489	105	0.054624065	0.046072839	0.84343874
57	436	112	0.0487042	0.049144362	1.09037442
58	65	21	0.007260947	0.009214568	1.269058629
59	62	19	0.006925827	0.00833699	1.203753769
60	69	23	0.007707775	0.010092146	1.309346204
61	62	23	0.006925827	0.010092146	1.457175615
62	65	22	0.007260947	0.009653357	1.329489992
63	58	15	0.006478999	0.006581834	1.015872055
64	64	24	0.00714924	0.010530935	1.47301448
65	63	29	0.007037534	0.012724879	1.808144759
66	63	23	0.007037534	0.010092146	1.434045843
67	62	26	0.006925827	0.011408513	1.647241999
68	65	33	0.007260947	0.014480035	1.994234988
69	62	24	0.006925827	0.010530935	1.520531076
70	67	26	0.007484361	0.011408513	1.524313492
71	58	31	0.006478999	0.013602457	2.099468914
72	64	30	0.00714924	0.013163668	1.8412681
73	63	29	0.007037534	0.012724879	1.808144759
74	55	25	0.006143878	0.010969724	1.785472097
75	61	23	0.00681412	0.010092146	1.481063739
76	62	31	0.006925827	0.013602457	1.964019307
77	62	29	0.006925827	0.012724879	1.837308384
78	58	24	0.006478999	0.010530935	1.625395288
79	60	31	0.006702413	0.013602457	2.029486617
80	58	31	0.006478999	0.013602457	2.099468914
81	60	33	0.006702413	0.014480035	2.160421237
82	60	36	0.006702413	0.015796402	2.356823168
83	57	35	0.006367292	0.015357613	2.411953535
84	62	36	0.006925827	0.015796402	2.280796614
85	60	33	0.006702413	0.014480035	2.160421237
86	62	34	0.006925827	0.014918824	2.154085691
87	55	31	0.006143878	0.013602457	2.2139854
88	58	40	0.006478999	0.017551558	2.708992147
89	56	37	0.006255585	0.016235191	2.595311227
90	60	40	0.006702413	0.017551558	2.618692409
91	58	44	0.006478999	0.019306713	2.979891362
92	58	36	0.006478999	0.015796402	2.438092932
93	56	36	0.006255585	0.015796402	2.52516768
94	58	36	0.006478999	0.015796402	2.438092932
95	56	39	0.006255585	0.017112769	2.73559832
96	58	40	0.006478999	0.017551558	2.708992147
97	56	39	0.006255585	0.017112769	2.73559832
98	57	40	0.006367292	0.017551558	2.756518325
99	53	40	0.005920465	0.017551558	2.964557444
100	40	37	0.004468275	0.016235191	3.633435717
N	8952	2279			

Notes: See the paper for how each of the components were calculated.

Table A.83: Summary Statistics of Bayesian Components

	Variable	Mean	Std. Dev.	Min	Max	N
Full Sample	$Prob(S_{c,2001} ASVAB_i)$	0.649	0.399	0	1	710
	$\frac{Prob(ASVAB_i S_c)}{Prob(ASVAB_i)}$	1.038	0.802	0	3.633	710
	$Prob(S_{c,1997})$	0.743	0.308	0	1	710
Interior Sample	$Prob(S_{c,2001} ASVAB_i)$	0.582	0.399	0	1	420
	$\frac{Prob(ASVAB_i S_c)}{Prob(ASVAB_i)}$	0.992	0.786	0	3.633	420
	$Prob(S_{c,1997})$	0.625	0.268	0	1	420

Notes: The interior sample are those who reported between 1% and 99% in 1997. See the paper for how each were calculated using the full NLSY97 sample.

Table A.84: Bayesian Updating Estimates

	<u>Full Sample</u>	<u>Interior Sample</u>	<u>Full Sample</u>
	(1)	(2)	(3)
$\hat{\theta}$	0.503 (0.0156)	0.517 (0.0251)	0.398 (0.0178)
Expectation = 0 in 1997			0.298 (0.0790)
Expectation = 100 in 1997			0.331 (0.0310)
Observations	710	420	710
Adjusted R^2	0.592	0.549	0.641

Notes: Significance stars are for whether or not θ is = 1

-Robust standard errors in parentheses

– * $p < 0.05$, ** $p < 0.01$, ***

Table A.85: 2001 Expectation Formation - Full Sample

	Y = Expectation of completing a college degree in 2001				
	(1)	(2)	(3)	(4)	(5)
College Expectations 1997	0.551*** (0.0374)	0.717*** (0.0263)	0.607*** (0.0258)	0.819*** (0.0172)	0.469*** (0.0383)
ASVAB Percentile	0.484*** (0.0517)				0.275*** (0.0567)
Took the ACT or SAT		18.16*** (3.020)			2.540 (3.133)
Enrolled in a 2yr College			32.63*** (3.942)		29.69*** (4.429)
Enrolled in a 4yr College			42.03*** (2.585)		33.41*** (3.264)
Change As				19.81 (11.38)	19.28 (15.26)
Change As & Bs				9.593 (11.42)	14.53 (15.28)
Change Bs				10.15 (11.63)	15.18 (15.43)
Change Bs & Cs				11.63 (11.49)	13.48 (15.39)
Change_Cs				17.39 (11.55)	21.45 (15.74)
Change Cs & Ds				17.35 (12.09)	18.82 (15.95)
Change Ds				15.40 (15.05)	20.60 (18.20)
Change Ds & other				19.65 (12.40)	22.80 (15.29)
Observations	577	710	710	710	577
Adjusted R^2	0.796	0.764	0.805	0.753	0.828

Notes: All of those who reported between 0% and 100% in in the initial round were used to obtain these estimates. When the ASVAB percentile is used only those without imputed ASVAB scores were utilized.

- Robust standard errors in parenthesis

—* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.86: 2001 Expectation Formation - 1 to 99

	Y = Expectation of completing a college degree in 2001				
	(1)	(2)	(3)	(4)	(5)
College Expectations 1997	0.540*** (0.0588)	0.794*** (0.0377)	0.668*** (0.0358)	0.851*** (0.0284)	0.480*** (0.0609)
ASVAB Percentile	0.507*** (0.0723)				0.334*** (0.0757)
Took the ACT or SAT		11.95** (4.160)			-1.349 (4.343)
Enrolled in a 2yr College			33.47*** (5.702)		31.06*** (6.158)
Enrolled in a 4yr College			41.53*** (3.332)		33.17*** (4.277)
Change As				26.97* (11.76)	15.70 (12.08)
Change As & Bs				15.13 (12.01)	9.934 (12.25)
Change Bs				12.39 (12.45)	6.956 (12.45)
Change Bs & Cs				20.43 (12.19)	9.666 (12.32)
Change Cs				23.03 (12.63)	15.13 (13.00)
Change Cs & Ds				16.49 (12.92)	7.314 (13.25)
Change Ds				24.59 (15.39)	19.41 (15.36)
Change Ds & other				31.34** (11.84)	27.11* (12.20)
Observations	341	420	420	420	341
Adjusted R^2	0.748	0.701	0.750	0.698	0.784

Notes: Only those who reported between 1% and 99% in in the initial round were used to obtain these estimates. When the ASVAB percentile is used only those without imputed ASVAB scores were utilized.

- Robust standard errors in parenthesis

—* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Chapter 3 Appendix

Table A.87: Full and Sub-sample Summary Statistics

Variable	Full Sample	1997 Sample	2001 Sample
College Expectations 1997	n/a	72.8%	n/a
College Expectations 2001	n/a	n/a	67.2%
Complete 4yr Degree	0.251	0.252	0.271
Representative Sample	0.751	0.747	0.753
Age in 1997	14.307	15.788	14.231
Age in 2001	19.055	20.561	18.948
Age in 2013	30.875	32.358	30.83
Male	0.512	0.504	0.511
White	0.519	0.52	0.511
–White Male	0.269	0.263	0.265
–White Female	0.251	0.256	0.246
Black	0.269	0.273	0.261
–Black Male	0.135	0.132	0.13
–Black Female	0.135	0.141	0.131
Hispanic	0.212	0.207	0.228
–Hispanic Male	0.109	0.109	0.116
–Hispanic Female	0.103	0.099	0.113
Household Size	4.549	4.492	4.583
No Parents Present	0.062	0.068	0.061
One Parent Household	0.315	0.324	0.301
Two Parent Household	0.623	0.609	0.638
Both Parents Less HS	0.099	0.095	0.104
Both Parents HS Degree	0.25	0.254	0.221
One Parent Some College	0.317	0.315	0.336
Both Parents Some College	0.061	0.056	0.07
One Parent College Degree	0.244	.241	0.243
Both Parents College Degree	0.097	0.093	0.101
ASVAB Percentile	44.762	44.858	45.435
–Imputed ASVAB	0.21	0.216	0.192
Gross HH Income	\$ 45312.23	\$ 46192.59	\$ 45276.61
–Imputed HH Income	0.267	0.275	0.255
High School GPA	2.818	2.797	2.832
Rural	0.268	0.269	0.242
Urban	0.732	0.731	0.758
No MSA	0.176	0.177	0.189
MSA Not Central City	0.502	0.499	0.495
MSA Central City	0.322	0.324	0.314
Census North East	0.176	0.173	0.174
Census North Central	0.228	0.238	0.214
Census South	0.374	0.38	0.381
Census West	0.222	0.21	0.232
Observations	8,984	3,511	1,946

Notes: The NLSY97 is composed of a representative portion and an over-sample portion. The percentage of the respondents in the representative sample is reported in the table above. The age in 2013 is reported because it is when the majority of the sample reported their highest educational achievement. For the full and 1997 sub-sample, the location variables are what is reported in 1997; the 2001 reported location is used for the 2001 sub-sample. See the paper for notes on the imputation of both the ASVAB percentile and Household Income.

Figure A.7: Frequency Distribution of College Completion Expectations

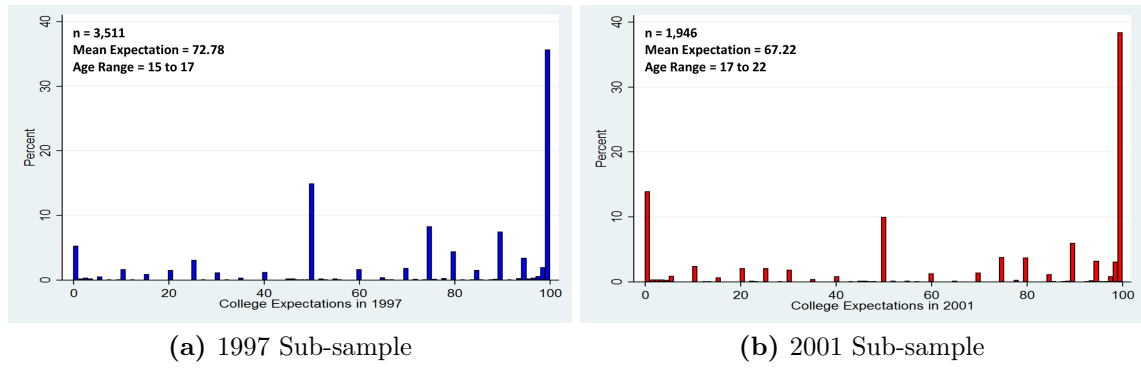


Table A.88: Four Year Degree Expectations and Attainment

	Percent Chance Have a 4 Year Degree by 30	Number of Students	Number that Attempt Degree	Number that Complete Degree	Percent that Attempt Degree	Percent that Complete Degree	Percent that Start and Complete Degree
1997 Sub-sample	0%	185	26	5	14.0%	2.7%	19.2%
	1 - 25 %	288	61	5	21.2%	1.7%	8.2%
	26 - 49 %	106	36	3	34%	2.8%	8.3%
	50 %	523	170	37	32.5%	7.1%	21.8%
	51 - 75 %	441	251	93	56.9%	21.1%	37.1%
	76 - 90 %	488	327	149	67.0%	30.5%	45.6%
	91 - 99 %	230	179	113	77.8%	49.1%	63.1%
	100 %	1,250	879	480	70.3%	38.4%	54.6%
	Total	3,511	1,929	885	54.9%	25.2%	45.9%
2001 Sub-sample	0 %	270	36	3	13.3%	1.1%	8.3%
	1 - 25 %	182	44	1	24.2%	0.5%	2.3%
	26 - 49 %	65	15	1	23.1%	1.5%	6.7%
	50 %	193	58	6	30.1%	3.1%	10.3 %
	51 - 75 %	131	75	18	57.3%	13.7%	24%
	76 - 90 %	215	130	49	60.5%	22.8%	37.7%
	91 - 99 %	143	119	85	83.2%	59.4%	71.4%
	100 %	747	640	365	85.7%	48.9%	57.0%
	Total	1,946	1,117	528	57.4%	27.1%	57.4%

Notes: Number that Attempt Degree includes those who reported their highest grade completed as 1 year of college or more.

Table A.89: 1997 Expectations and Completion Summary Statistics

Variable	Completed College			Did Not Complete College		
	Full Sample	90-99%	100%	Full Sample	90-99%	100%
College Expectations 1997	91.00%	93.71%	100%	66.64%	92.56%	100%
Representative Sample	0.836	0.897	0.827	0.717	0.778	0.719
Age in 1997	15.78	15.77	15.79	15.79	15.77	15.75
Age in 2001	20.52	20.51	20.52	20.57	20.59	20.55
Age in 2013	32.38	32.37	32.38	32.35	32.38	32.30
Male	0.421	0.485	0.358	0.532	0.5	0.435
White	0.68	0.804	0.633	0.465	0.521	0.438
-White Male	0.298	0.407	0.229	0.251	0.267	0.169
-White Female	0.382	0.397	0.404	0.214	0.253	0.269
Black	0.202	0.113	0.254	0.297	0.219	0.377
-Black Male	0.077	0.039	0.096	0.151	0.076	0.179
-Black Female	0.125	0.074	0.158	0.147	0.142	0.197
Hispanic	0.118	0.083	0.113	0.237	0.26	0.186
-Hispanic Male	0.046	0.039	0.033	0.129	0.156	0.087
-Hispanic Female	0.071	0.044	0.079	0.108	0.104	0.099
Household Size	4.339	4.338	4.254	4.544	4.549	4.456
No Parents	0.034	0.015	0.046	0.079	0.069	0.079
One Parent	0.208	0.196	0.198	0.363	0.302	0.34
Two Parents	0.758	0.789	0.756	0.558	0.628	0.581
Both Parents Less HS	0.028	0.02	0.021	0.117	0.097	0.073
Both Parents HS Degree	0.163	0.103	0.154	0.284	0.288	0.253
One Parent Some College	0.386	0.353	0.423	0.291	0.292	0.351
Both Parents Some College	0.078	0.054	0.083	0.049	0.042	0.058
One Parent College Degree	0.501	0.623	0.494	0.153	0.233	0.216
Both Parents College Degree	0.25	0.333	0.246	0.04	0.08	0.064
ASVAB Percentile	64.41	73.51	62.88	38.27	46.58	40.75
-Imputed ASVAB	0.142	0.142	0.135	0.241	0.229	0.235
Gross HH Income	\$ 66,869.70	\$ 74,619.30	\$ 66,738.90	\$ 39,224.10	\$ 46,446.90	\$ 42,877.30
-Imputed HH Income	0.267	0.206	0.292	0.278	0.299	0.291
High School GPA	3.233	3.329	3.278	2.617	2.745	2.701
Mostly As 8th grade	0.371	0.426	0.41	0.065	0.094	0.118
As and Bs 8th grade	0.301	0.328	0.288	0.182	0.229	0.244
Mostly Bs 8th grade	0.148	0.132	0.158	0.14	0.153	0.164
Bs and Cs 8th grade	0.12	0.078	0.1	0.3	0.316	0.281
Mostly Cs 8th grade	0.042	0.025	0.029	0.157	0.118	0.104
Cs and Ds 8th grade	0.015	0	0.013	0.103	0.049	0.066
Mostly Ds 8th grade	0.001	0	0	0.031	0.021	0.016
Mostly Below Ds 8th grade	0	0	0	0.014	0.01	0.006
Rural	0.256	0.265	0.258	0.274	0.205	0.273
Urban	0.744	0.735	0.742	0.726	0.795	0.727
No MSA	0.142	0.127	0.14	0.188	0.142	0.196
MSA Not Central City	0.567	0.598	0.569	0.476	0.5	0.46
MSA Central City	0.29	0.275	0.292	0.335	0.358	0.344
Census North East	0.198	0.196	0.213	0.164	0.177	0.16
Census North Central	0.266	0.333	0.219	0.228	0.24	0.225
Census South	0.33	0.245	0.385	0.397	0.354	0.452
Census West	0.207	0.225	0.183	0.211	0.229	0.164
Observations	885	204	480	2,626	288	770

Notes:

Table A.90: 2001 Expectations and Completion Summary Statistics

Variable	Completed College			Did Not Complete College		
	Full Sample	90-99%	100%	Full Sample	90-99%	100%
College Expectations 2001	96.08%	95.22%	100	56.48%	92.72%	100%
Representative Sample	0.881	0.941	0.858	0.705	0.75	0.654
Age in 1997	14.33	14.01	14.52	14.20	13.89	14.14
Age in 2001	19.00	18.69	19.20	18.93	18.66	18.88
Age in 2013	30.91	30.61	31.09	30.80	30.52	30.71
Male	0.432	0.568	0.392	0.54	0.514	0.463
White	0.718	0.805	0.685	0.434	0.464	0.393
-White Male	0.318	0.458	0.268	0.245	0.243	0.186
-White Female	0.4	0.347	0.416	0.188	0.221	0.207
Black	0.163	0.093	0.186	0.298	0.271	0.356
-Black Male	0.059	0.042	0.068	0.157	0.15	0.165
-Black Female	0.104	0.051	0.118	0.141	0.121	0.191
Hispanic	0.119	0.102	0.129	0.269	0.264	0.251
-Hispanic Male	0.055	0.068	0.055	0.138	0.121	0.113
-Hispanic Female	0.064	0.034	0.074	0.13	0.143	0.139
Household Size	4.398	4.356	4.381	4.652	4.507	4.66
No Parents	0.036	0.034	0.041	0.07	0.071	0.065
One Parent	0.195	0.178	0.208	0.341	0.293	0.33
Two Parents	0.769	0.788	0.751	0.59	0.636	0.605
Both Parents Less HS	0.028	0.017	0.033	0.132	0.086	0.086
Both Parents HS Degree	0.148	0.136	0.145	0.248	0.221	0.225
One Parent Some College	0.411	0.331	0.433	0.308	0.357	0.393
Both Parents Some College	0.085	0.068	0.082	0.065	0.093	0.097
One Parent College Degree	0.521	0.593	0.504	0.139	0.214	0.225
Both Parents College Degree	0.267	0.347	0.244	0.039	0.064	0.071
ASVAB Percentile	65.97	72.69	64.61	37.79	43.84	44.37
-Imputed ASVAB	0.129	0.136	0.132	0.216	0.179	0.223
Gross HH Income	\$ 64,366.50	\$ 62,196.10	\$ 64,480.60	\$ 38,168.40	\$ 46,871.70	\$ 44,436.30
-Imputed HH Income	0.242	0.271	0.227	0.26	0.329	0.301
Highest Grade Completed 2001	12.515	12.161	12.748	11.251	11.457	11.835
Enrolled in 2yr College	0.098	0.102	0.099	0.086	0.152	0.144
Enrolled in 4yr College	0.555	0.441	0.638	0.057	0.065	0.168
High School GPA	3.279	3.386	3.273	2.632	2.677	2.805
Mostly As HS	0.301	0.373	0.29	0.041	0.057	0.063
As and Bs HS	0.32	0.347	0.323	0.166	0.143	0.27
Mostly Bs HS	0.197	0.169	0.195	0.166	0.179	0.202
Bs and CsHS	0.114	0.085	0.118	0.281	0.293	0.249
Mostly Cs HS	0.036	0.017	0.041	0.172	0.193	0.131
Cs and Ds HS	0.009	0	0.011	0.095	0.071	0.037
Mostly DsHS	0	0	0	0.022	0.007	0.01
Mostly below Ds HS	0	0	0	0.016	0	0.005
Rural	0.237	0.212	0.233	0.26	0.207	0.191
Urban	0.763	0.788	0.767	0.74	0.793	0.809
No MSA	0.148	0.119	0.148	0.19	0.214	0.123
MSA Not Central City	0.563	0.551	0.564	0.476	0.507	0.5
MSA Central City	0.29	0.331	0.288	0.334	0.279	0.377
Census North East	0.216	0.161	0.241	0.159	0.171	0.165
Census North Central	0.263	0.246	0.252	0.206	0.229	0.209
Census South	0.301	0.28	0.321	0.402	0.35	0.414
Census West	0.22	0.314	0.186	0.233	0.25	0.212
Observations	528	118	365	1,418	140	382

Notes:

Table A.91: Probit Model Predictions of College Completion

Controls Used to Predict Completion	Sample	N	Mean	Std. Deviation	Min	Max	Percent that Complete Degree by 2013
1997 Grades	Full Sample:	8865	0.254	0.255	0.000	0.976	25.14%
	Do Not Complete College	6606	0.167	0.185	0.000	0.976	0%
	Complete College	2259	0.510	0.260	0.006	0.976	100%
	1997 Sub-sample:	3475	0.257	0.255	0.000	0.976	25.21%
	Do Not Complete College	2590	0.170	0.185	0.000	0.949	0%
	Complete College	885	0.512	0.263	0.011	0.976	100%
2001 Enrollment	Full Sample:	7808	0.269	0.296	0.001	0.987	27.07%
	Do Not Complete College	5694	0.148	0.192	0.001	0.978	0%
	Complete College	2114	0.593	0.283	0.010	0.987	100%
	2001 Sub-sample:	1933	0.275	0.298	0.001	0.983	27.13%
	Do Not Complete College	1408	0.152	0.197	0.001	0.978	0%
	Complete College	525	0.606	0.270	0.013	0.983	100%

Notes: The mean reported is the average predicted probability of completing college for the sample utilized from a Probit model that included similar controls to essay 1 but excluded the student's expectations. The two portions represent different controls used in each model as there were concerns about selection based on who chose to report their High School GPA.

Table A.92: Probit Model Predictions of College Completion - 1997 Sub-sample

Controls Used to Predict Completion	Sample	N	Mean	Std. Deviation	Min	Max	Percent that Complete Degree by 2013
1997 Grades	Full Sample:	8865	0.254	0.255	0.000	0.976	25.14%
	Do Not Complete College	6606	0.167	0.185	0.000	0.976	0%
	Complete College	2259	0.510	0.260	0.006	0.976	100%
	1997 Sub-sample:	3475	0.257	0.255	0.000	0.976	25.21%
	Do Not Complete College	2590	0.170	0.185	0.000	0.949	0%
	Complete College	885	0.512	0.263	0.011	0.976	100%
1997 HS GPA	Full Sample:	6004	0.296	0.273	0.000	0.986	29.58%
	Do Not Complete College	4228	0.189	0.197	0.000	0.981	0%
	Complete College	1776	0.550	0.259	0.003	0.986	100%
	1997 Sub-sample:	2561	0.295	0.272	0.000	0.986	29.29%
	Do Not Complete College	1811	0.191	0.196	0.000	0.966	0%
	Complete College	750	0.548	0.264	0.006	0.986	100%

Notes: The mean reported is the average predicted probability of completing college for the sample utilized from a Probit model that included similar controls to essay 1 but excluded the student's expectations. The two portions represent different controls used in each model as there were concerns about selection based on who chose to report their High School GPA.

Table A.93: Probit Model Predictions of College Completion - 2001 Sub-sample

Controls Used to Predict Completion	Sample	N	Mean	Std. Deviation	Min	Max	Percent that Complete Degree by 2013
2001 HS GPA	Full Sample:	5518	0.307	0.318	0.000	0.994	30.86%
	Do Not Complete College	3815	0.160	0.205	0.000	0.988	0%
	Complete College	1703	0.636	0.275	0.002	0.994	100%
	2001 Sub-sample:	1348	0.316	0.320	0.000	0.994	30.95%
	Do Not Complete College	930	0.165	0.214	0.000	0.988	0%
	Complete College	418	0.651	0.257	0.047	0.994	100%
2001 Enrollment	Full Sample:	7808	0.269	0.296	0.001	0.987	27.07%
	Do Not Complete College	5694	0.148	0.192	0.001	0.978	0%
	Complete College	2114	0.593	0.283	0.010	0.987	100%
	2001 Sub-sample:	1933	0.275	0.298	0.001	0.983	27.13%
	Do Not Complete College	1408	0.152	0.197	0.001	0.978	0%
	Complete College	525	0.606	0.270	0.013	0.983	100%

Notes: The mean reported is the average predicted probability of completing college for the sample utilized from a Probit model that included similar controls to essay 1 but excluded the student's expectations. The two portions represent different controls used in each model as there were concerns about selection based on who chose to report their High School GPA.

Figure A.8: Primary Alignment Measure for each Sub-sample

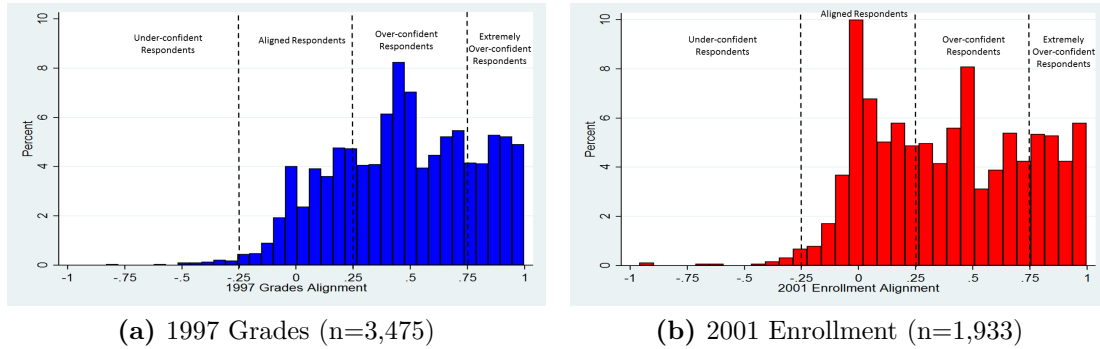


Figure A.9: Primary Alignment Measures using the Overlapping Sample

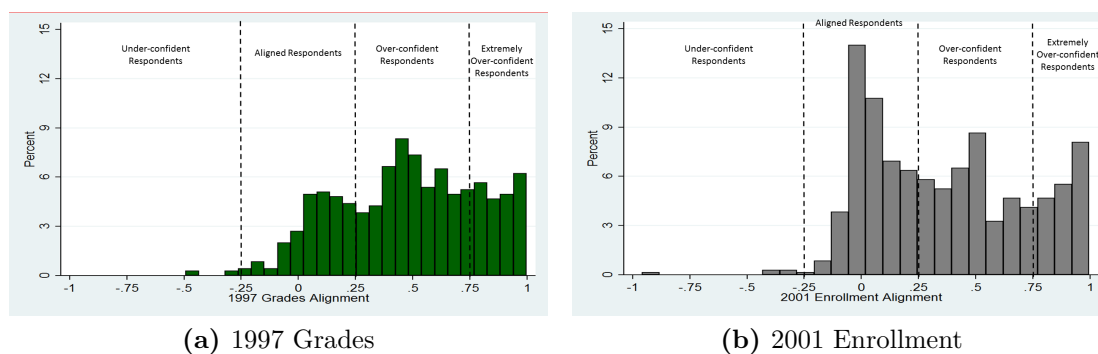


Table A.94: Summary Statistics for the Secondary Alignment Measure

Sample	N	Mean	Standard Deviation	Min	Max	Percent that Complete Degree by 2013
1997 Sub-Sample						
Full:	3406	8.87	13.65	0	98.10	25.98%
-Under-confident	232	0.10	0.21	0	0.74	3.88%
-Aligned	288	1.09	0.11	0.77	1.25	68.75%
-Over-confident	576	1.58	0.21	1.25	2.00	52.95%
-Extremely Over-confident	2310	12.54	15.26	2.00	98.10	16.15%
2001 Sub-Sample						
Full:	1898	8.43	14.85	0	97.52	27.66%
-Under-confident	301	0.04	0.14	0	0.75	1.66%
-Aligned	236	1.10	0.09	0.75	1.25	72.46%
-Over-confident	288	1.55	0.20	1.25	2.00	62.85%
-Extremely Over-confident	1073	14.25	17.66	2.00	97.52	15.66%

Notes: The secondary measure of alignment is the ration of the individual's reported expectations and the estimated probability that they will complete a college degree using the sample specific controls outlined in the paper. Those who received an alignment measure of over 100 are not included in these statistics.

Table A.95: 1997 Expectation Alignment Summary Statistics

	1997 Grades Alignment					1997 GPA Alignment				
	N	Mean	Std. Dev	Min	Max	N	Mean	Std. Dev	Min	Max
Male	1743	0.481	0.310	-0.492	0.999	1265	0.471	0.314	-0.481	1.000
Female	1732	0.465	0.310	-0.829	0.994	1296	0.451	0.316	-0.813	1.000
White	1808	0.409	0.298	-0.829	0.997	1433	0.396	0.306	-0.813	0.999
Black	950	0.570	0.311	-0.454	0.999	641	0.556	0.315	-0.396	1.000
Hispanic	717	0.507	0.299	-0.365	0.994	487	0.525	0.296	-0.425	1.000
White Male	911	0.410	0.298	-0.492	0.997	718	0.398	0.305	-0.481	0.999
White Female	897	0.409	0.298	-0.829	0.993	715	0.395	0.308	-0.813	0.999
Black Male	459	0.594	0.308	-0.454	0.999	287	0.595	0.309	-0.396	1.000
Black Female	491	0.547	0.313	-0.422	0.994	354	0.524	0.318	-0.272	1.000
Hispanic Male	373	0.517	0.293	-0.164	0.994	260	0.534	0.286	-0.170	0.999
Hispanic Female	344	0.496	0.306	-0.365	0.994	227	0.515	0.307	-0.425	1.000

Notes: The alignment measures are difference between the percent chance of completing a college degree by 30 reported by the individual and the predicted probability that they would complete the degree based on a set of standard covariates. For the full list of covariates used, see the paper.

Table A.96: 2001 Expectation Alignment Summary Statistics

	2001 Enrollment Alignment					2001 GPA Alignment				
	N	Mean	Std. Dev	Min	Max	N	Mean	Std. Dev	Min	Max
Male	985	0.401	0.348	-0.957	0.998	663	0.407175	0.349	-0.955	1.000
Female	948	0.392	0.341	-0.925	0.997	685	0.380151	0.337	-0.882	0.995
White	987	0.328	0.322	-0.957	0.994	732	0.318356	0.321	-0.955	0.997
Black	505	0.490	0.353	-0.663	0.998	343	0.481372	0.348	-0.508	1.000
Hispanic	441	0.446	0.352	-0.435	0.995	273	0.484298	0.346	-0.354	0.994
White Male	511	0.339	0.325	-0.957	0.994	378	0.336831	0.327	-0.955	0.997
White Female	476	0.316	0.318	-0.925	0.992	354	0.298628	0.313	-0.882	0.975
Black Male	251	0.496	0.366	-0.663	0.998	159	0.513804	0.354	-0.195	1.000
Black Female	254	0.484	0.341	-0.246	0.997	184	0.453346	0.342	-0.508	0.995
Hispanic Male	223	0.439	0.353	-0.153	0.995	126	0.483652	0.357	-0.277	0.994
Hispanic Female	218	0.453	0.353	-0.435	0.984	147	0.484852	0.337	-0.354	0.989

Notes: The alignment measures are difference between the percent chance of completing a college degree by 30 reported by the individual and the predicted probability that they would complete the degree based on a set of standard covariates. For the full list of covariates used, see the paper.

Table A.97: Primary Alignment Measures for the 1997 and 2001 Sub-samples

	1997 Grades Alignment					2001 GPA Alignment				
	N	Mean	Std. Dev	Min	Max	N	Mean	Std. Dev	Min	Max
Male	1743	0.481	0.310	-0.492	0.999	985	0.401	0.348	-0.957	0.998
Female	1732	0.465	0.310	-0.829	0.994	948	0.392	0.341	-0.925	0.997
White	1808	0.409	0.298	-0.829	0.997	987	0.328	0.322	-0.957	0.994
Black	950	0.570	0.311	-0.454	0.999	505	0.490	0.353	-0.663	0.998
Hispanic	717	0.507	0.299	-0.365	0.994	441	0.446	0.352	-0.435	0.995
White Male	911	0.410	0.298	-0.492	0.997	511	0.339	0.325	-0.957	0.994
White Female	897	0.409	0.298	-0.829	0.993	476	0.316	0.318	-0.925	0.992
Black Male	459	0.594	0.308	-0.454	0.999	251	0.496	0.366	-0.663	0.998
Black Female	491	0.547	0.313	-0.422	0.994	254	0.484	0.341	-0.246	0.997
Hispanic Male	373	0.517	0.293	-0.164	0.994	223	0.439	0.353	-0.153	0.995
Hispanic Female	344	0.496	0.306	-0.365	0.994	218	0.453	0.353	-0.435	0.984

Notes: The alignment measures are difference between the percent chance of completing a college degree by 30 reported by the individual and the predicted probability that they would complete the degree based on a set of standard covariates. For the full list of covariates used, see the paper.

Table A.98: 1997 Expectation Alignment Summary Statistics - Overlapping Sample

	1997 Grades Alignment					1997 GPA Alignment				
	N	Mean	Std. Dev	Min	Max	N	Mean	Std. Dev	Min	Max
Male	343	0.484	0.307	-0.492	0.998	254	0.458	0.317	-0.396	0.999
Female	365	0.477	0.306	-0.313	0.988	285	0.465	0.312	-0.318	0.998
White	358	0.408	0.289	-0.492	0.997	289	0.396	0.305	-0.383	0.999
Black	182	0.585	0.308	-0.454	0.998	134	0.557	0.308	-0.396	0.996
Hispanic	168	0.521	0.301	-0.155	0.989	116	0.514	0.308	-0.110	0.998
White Male	179	0.414	0.292	-0.492	0.997	144	0.401	0.312	-0.383	0.999
White Female	179	0.402	0.287	-0.313	0.988	145	0.392	0.298	-0.318	0.998
Black Male	81	0.625	0.290	-0.454	0.998	57	0.574	0.302	-0.396	0.984
Black Female	101	0.552	0.319	-0.217	0.988	77	0.545	0.314	-0.160	0.996
Hispanic Male	83	0.495	0.309	-0.075	0.989	53	0.488	0.314	-0.087	0.962
Hispanic Female	85	0.545	0.293	-0.155	0.973	63	0.537	0.304	-0.110	0.998

Notes: The alignment measures are difference between the percent chance of completing a college degree by 30 reported by the individual and the predicted probability that they would complete the degree based on a set of standard covariates. For the full list of covariates used, see the paper.

Table A.99: 2001 Expectation Alignment Summary Statistics - Overlapping Sample

	2001 Enrollment Alignment					2001 GPA Alignment				
	N	Mean	Std. Dev	Min	Max	N	Mean	Std. Dev	Min	Max
Male	345	0.367	0.345	-0.957	0.997	250	0.367	0.358	-0.955	0.997
Female	362	0.369	0.346	-0.404	0.997	282	0.374	0.342	-0.354	0.995
White	358	0.285	0.314	-0.957	0.994	286	0.297	0.322	-0.955	0.997
Black	182	0.472	0.349	-0.121	0.997	132	0.464	0.358	-0.195	0.995
Hispanic	167	0.430	0.363	-0.404	0.989	114	0.447	0.366	-0.354	0.989
White Male	180	0.306	0.328	-0.957	0.994	142	0.309	0.342	-0.955	0.997
White Female	178	0.265	0.299	-0.359	0.956	144	0.285	0.303	-0.302	0.975
Black Male	82	0.443	0.349	-0.077	0.997	56	0.462	0.365	-0.195	0.993
Black Female	100	0.497	0.350	-0.121	0.997	76	0.466	0.355	-0.191	0.995
Hispanic Male	83	0.424	0.356	-0.146	0.989	52	0.423	0.372	-0.273	0.981
Hispanic Female	84	0.437	0.371	-0.404	0.984	62	0.467	0.363	-0.354	0.989

Notes: The alignment measures are difference between the percent chance of completing a college degree by 30 reported by the individual and the predicted probability that they would complete the degree based on a set of standard covariates. For the full list of covariates used, see the paper.

Table A.100: 1997 Expectations and Alignment

Variable	Full Sample	Under-confident	Aligned	Over-confident	Extremely Over-confident
College Expectation 1997	0.728	0.078	0.459	0.765	0.976
Age in 1997	15.788	16	15.848	15.774	15.755
Representative Sample	0.747	0.846	0.804	0.748	0.679
Male	0.504	0.5	0.493	0.505	0.508
White	0.52	0.731	0.623	0.546	0.336
-White male	0.263	0.385	0.315	0.274	0.167
-White female	0.256	0.346	0.308	0.272	0.169
Black	0.273	0.231	0.187	0.25	0.425
-Black male	0.132	0.115	0.083	0.119	0.22
-Black female	0.141	0.115	0.104	0.131	0.205
Hispanic	0.207	0.038	0.19	0.204	0.239
-Hispanic male	0.109	0	0.094	0.112	0.122
-Hispanic female	0.099	0.038	0.095	0.092	0.117
No Parents	0.068	0	0.059	0.053	0.109
One Parent	0.324	0.231	0.266	0.324	0.388
Two Parents	0.609	0.769	0.675	0.623	0.503
Both Parents Less HS	0.095	0.038	0.104	0.083	0.113
Both Parents HS Degree	0.254	0.192	0.219	0.247	0.317
One Parent Some College	0.315	0.5	0.256	0.37	0.255
Both Parents Some College	0.056	0.038	0.052	0.071	0.033
One Parent College Degree	0.241	0.385	0.363	0.231	0.107
Both Parents College Degree	0.093	0.154	0.217	0.063	0.014
ASVAB Percentile	44.858	62.947	53.707	46.565	30.747
-Imputed ASVAB	0.216	0.231	0.192	0.201	0.282
Gross HH Income	46192.6	54537.6	57104.8	46135.5	33480.5
-Imputed HH Income	0.275	0.269	0.248	0.281	0.293
High School GPA	2.797	2.935	2.966	2.817	2.541
Mostly As 8th grade	0.142	0.346	0.278	0.131	0.009
As and Bs 8th grade	0.212	0.385	0.201	0.257	0.123
Mostly Bs 8th grade	0.142	0.115	0.104	0.155	0.164
Bs and Cs 8th grade	0.255	0.154	0.182	0.235	0.404
Mostly Cs 8th grade	0.128	0	0.105	0.124	0.173
Cs and Ds 8th grade	0.081	0	0.099	0.068	0.096
Mostly Ds 8th grade	0.024	0	0.022	0.023	0.026
Mostly Below Ds 8th grade	0.01	0	0	0	0
Mostly As HS	0.104	0.115	0.204	0.091	0.018
As and Bs HS	0.187	0.269	0.176	0.208	0.145
Mostly Bs HS	0.154	0.192	0.134	0.165	0.158
Bs and Cs HS	0.254	0.269	0.195	0.256	0.326
Mostly Cs HS	0.13	0	0.103	0.121	0.184
Cs and Ds HS	0.076	0	0.083	0.071	0.085
Mostly Ds HS	0.021	0	0.017	0.02	0.021
Mostly below Ds HS	0.016	0	0.02	0.011	0.013
Rural	0.269	0.269	0.3	0.263	0.257
Urban	0.731	0.731	0.7	0.737	0.743
No MSA	0.177	0.154	0.154	0.187	0.19
MSA Not Central City	0.499	0.5	0.559	0.507	0.416
MSA Central City	0.324	0.346	0.287	0.306	0.394
Census North East	0.173	0.038	0.175	0.185	0.148
Census North Central	0.238	0.269	0.266	0.24	0.202
Census South	0.38	0.385	0.343	0.358	0.466
Census West	0.21	0.308	0.215	0.217	0.185

Notes: The alignment measures are difference between the percent chance of completing a college degree by 30 reported by the individual and the predicted probability that they would complete the degree based on a set of standard covariates. For the full list of covariates used, see the paper. The under-confident are those with an alignment measure between -.1 and -.24. The aligned are those who score between -.25 and .25. The over-confident are those who score between .26 and .75, and the extremely over-confident score greater than .75.

Table A.101: 2001 Expectations and Alignment

Variable	Full Sample	Under-confident	Aligned	Over-confident	Extremely Over-confident
College Expectation 2001	0.672	0.071	0.403	0.787	0.969
Age in 2001	18.948	18.583	19.203	18.752	18.835
Representative Sample	0.753	0.875	0.818	0.744	0.64
Male	0.511	0.5	0.509	0.501	0.517
White	0.511	0.792	0.605	0.513	0.323
–White male	0.265	0.458	0.299	0.269	0.182
–White female	0.246	0.333	0.306	0.244	0.14
Black	0.261	0.083	0.206	0.251	0.397
–Black male	0.13	0.042	0.106	0.116	0.2
–Black female	0.131	0.042	0.1	0.135	0.197
Hispanic	0.228	0.125	0.189	0.236	0.281
–Hispanic male	0.116	0	0.104	0.116	0.135
–Hispanic female	0.113	0.125	0.085	0.12	0.145
No Parents	0.061	0	0.045	0.06	0.091
One Parent	0.301	0.208	0.285	0.296	0.357
Two Parents	0.638	0.792	0.669	0.644	0.552
Both Parents Less HS	0.104	0	0.112	0.095	0.111
Both Parents HS Degree	0.221	0.208	0.196	0.231	0.261
One Parent Some College	0.336	0.5	0.309	0.373	0.298
Both Parents Some College	0.07	0.083	0.053	0.083	0.074
One Parent College Degree	0.243	0.417	0.316	0.231	0.118
Both Parents College Degree	0.101	0.125	0.17	0.075	0.027
ASVAB Percentile	45.435	71.654	49.486	47.652	32.847
–Imputed ASVAB	0.192	0.167	0.169	0.183	0.249
Gross HH Income	45276.6	55361.9	50465.2	45842.9	34567.2
–Imputed HH Income	0.255	0.292	0.233	0.257	0.293
Enrolled in HS	0.289	0.417	0.174	0.376	0.335
Enrolled in 2yr College	0.09	0.25	0.041	0.123	0.106
Enrolled in 4yr College	0.192	0.208	0.311	0.179	0
High School GPA	2.832	2.789	2.891	2.869	2.674
Mostly As 8th grade	0.155	0.25	0.184	0.179	0.062
As and Bs 8th grade	0.214	0.25	0.2	0.221	0.217
Mostly Bs 8th grade	0.136	0.167	0.122	0.144	0.148
Bs and Cs 8th grade	0.223	0.25	0.213	0.205	0.273
Mostly Cs 8th grade	0.142	0.083	0.126	0.137	0.18
Cs and Ds 8th grade	0.076	0	0.085	0.068	0.069
Mostly Ds 8th grade	0.027	0	0.036	0.023	0.025
Mostly Below Ds 8th grade	0.015	0	0.019	0.008	0.02
Mostly As HS	0.112	0.083	0.15	0.116	0.037
As and Bs HS	0.208	0.417	0.17	0.229	0.227
Mostly Bs HS	0.174	0.125	0.163	0.204	0.145
Bs and Cs HS	0.235	0.167	0.221	0.213	0.303
Mostly Cs HS	0.135	0.125	0.14	0.113	0.172
Cs and Ds HS	0.072	0.042	0.08	0.067	0.064
Mostly Ds HS	0.016	0	0.016	0.016	0.012
Mostly below Ds HS	0.011	0	0.015	0.011	0.002
Rural	0.242	0.292	0.294	0.22	0.192
Urban	0.758	0.708	0.706	0.78	0.808
No MSA	0.189	0.208	0.217	0.18	0.167
MSA Not Central City	0.495	0.667	0.528	0.473	0.463
MSA Central City	0.314	0.125	0.255	0.347	0.369
Census North East	0.174	0.208	0.189	0.161	0.165
Census North Central	0.214	0.208	0.209	0.223	0.207
Census South	0.381	0.25	0.376	0.376	0.414
Census West	0.232	0.333	0.226	0.24	0.214

Notes: The alignment measures are difference between the percent chance of completing a college degree by 30 reported by the individual and the predicted probability that they would complete the degree based on a set of standard covariates. For the full list of covariates used, see the paper. The under-confident are those with an alignment measure between -.1 and -.24. The aligned are those who score between -.25 and .25. The over-confident are those who score between .26 and .75, and the extremely over-confident score greater than .75.

Table A.102: Estimates of Expectation Alignment in 1997

	$Y = (Expectation_{i,1997} - P(CompleteCollege_{i,2013}))^2$	
	1997 Sub-sample	Overlapping 700
Age in 1997	-0.0173** (0.00632)	-0.0169 (0.0144)
Male	-0.00223 (0.00912)	0.00296 (0.0205)
White	-0.0591*** (0.0149)	-0.0703* (0.0351)
Hispanic	-0.0423** (0.0156)	-0.0414 (0.0344)
One Parent HH	-0.0704** (0.0224)	-0.127* (0.0530)
Two Parent HH	-0.0563** (0.0211)	-0.0954 (0.0490)
Both Parents Less than HS	-0.0628** (0.0212)	-0.0344 (0.0469)
Both Parents HS Degree	-0.0221 (0.0152)	-0.0257 (0.0343)
One Parent Some College	-0.00290 (0.0133)	-0.0196 (0.0282)
Both Parents Some College	-0.0509** (0.0171)	0.000296 (0.0333)
One Parent College Degree	-0.0396** (0.0129)	-0.0358 (0.0276)
Both Parents College Degree	-0.0956*** (0.0141)	-0.0901** (0.0316)
ASVAB Percentile	-0.00209*** (0.000213)	-0.00230*** (0.000483)
Gross HH Income Quantile	-0.0139** (0.00478)	-0.0120 (0.00996)
As 8th Grade	-0.156*** (0.0172)	-0.117** (0.0361)
As & Bs 8th Grade	-0.0630*** (0.0170)	-0.0165 (0.0357)
Bs 8th Grade	0.00210 (0.0186)	0.0475 (0.0393)
Bs & Cs 8th Grade	0.0328 (0.0180)	0.0626 (0.0391)
Cs & Ds 8th Grade	-0.0372 (0.0253)	-0.0310 (0.0599)
Ds 8th Grade	-0.0346 (0.0388)	0.0378 (0.0756)
Enrolled in HS in 1997	0.163*** (0.0243)	0.130* (0.0588)
Urban in 1997	0.0191 (0.0118)	-0.00491 (0.0263)
In MSA, Not Central City in 1997	-0.0313* (0.0135)	0.00984 (0.0303)
In Central City MSA in 1997	-0.0186 (0.0161)	0.0356 (0.0351)
North East Census in 1997	-0.0126 (0.0129)	-0.0169 (0.0284)
North Central Census in 1997	-0.00698 (0.0117)	-0.00172 (0.0271)
West Census in 1997	-0.0142 (0.0129)	-0.0624* (0.0281)
Observations	3475	708
Adjusted R^2	0.215	0.227

Notes: The model included a statistically significant constant with a coefficient estimate of approximately .75 but was not included. Each of the specifications above was also estimated without a constant and the majority of the estimates excluding age were similar in sign and significance.

- Robust standard errors in parentheses

- * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.103: Estimates of Expectation Alignment in 2001

	$Y = (Expectation_{i,2001} - P(CompleteCollege_{i,2013}))^2$	
	2001 Sub-sample	Overlapping 700
Age in 2001	-0.00214 (0.00612)	-0.00124 (0.0179)
Male	0.0171 (0.0129)	-0.00514 (0.0220)
White	-0.0619** (0.0214)	-0.0697 (0.0386)
Hispanic	-0.0324 (0.0225)	-0.0493 (0.0385)
One Parent HH	-0.0452 (0.0325)	0.0226 (0.0511)
Two Parent HH	-0.0331 (0.0311)	0.0472 (0.0459)
Both Parents Less than HS	-0.0698* (0.0278)	-0.0772 (0.0467)
Both Parents HS Degree	-0.0251 (0.0219)	-0.0537 (0.0359)
One Parent Some College	-0.00292 (0.0190)	-0.0524 (0.0319)
Both Parents Some College	0.00742 (0.0247)	0.0177 (0.0453)
One Parent College Degree	-0.0134 (0.0189)	-0.00709 (0.0319)
Both Parents College Degree	-0.0851*** (0.0225)	-0.0957* (0.0393)
ASVAB Percentile	-0.00116*** (0.000309)	-0.000110 (0.000556)
Gross HH Income Quantile	-0.00617 (0.00656)	-0.00907 (0.0111)
As HS	0.0284 (0.0249)	0.0132 (0.0424)
As & Bs HS	0.0797*** (0.0232)	0.107** (0.0411)
Bs HS	0.0331 (0.0237)	0.0498 (0.0402)
Bs & Cs HS	0.0227 (0.0227)	0.0442 (0.0377)
Cs & Ds HS	-0.0595* (0.0300)	-0.0203 (0.0524)
Ds HS	-0.123* (0.0586)	0.00161 (0.116)
Below Ds in HS	-0.209*** (0.0452)	-0.161** (0.0491)
Enrolled in HS in 2001	0.00544 (0.0217)	0.0345 (0.145)
Enrolled in a 2yr college in 2001	0.0458 (0.0235)	0.0507 (0.0336)
Enrolled in a 4yr college in 2001	-0.200*** (0.0157)	-0.199*** (0.0230)
Urban in 2001	0.0421* (0.0169)	0.0619* (0.0302)
In MSA, Not Central City in 2001	0.0193 (0.0181)	0.0252 (0.0313)
In Central City MSA in 2001	0.0300 (0.0216)	0.00457 (0.0366)
North East Census in 2001	0.0410* (0.0190)	0.00906 (0.0332)
North Central Census in 2001	0.0254 (0.0168)	0.00914 (0.0291)
West Census in 2001	-0.0169 (0.0180)	-0.0305 (0.0318)
Observations	1933	707
Adjusted R^2	0.171	0.161

Notes: The model included a statistically significant constant with a coefficient estimate of approximately .44 for the full model but was not included; the constant was not statistically significant for the overlapping 700. Each of the specifications above was also estimated without a constant and the majority of the estimates excluding age and enrollment in high school or a two-year college were similar in sign and significance.

- Robust standard errors in parentheses

- * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.104: Expectation Alignment and Completion

1997 Sub-Sample				2001 Sub-sample		
	Expectations	N	% that Complete College Degree	Expectations	N	% that Complete College Degree
Aligned Respondents	0%	164	1.22%	0%	252	0.79%
	1-25%	271	1.48%	1-25%	177	0.56%
	26-49%	35	2.94%	26-49%	22	4.55%
	50%	81	23.46%	50%	18	16.67%
	51-74%	22	54.55%	51-74%	4	25.00%
	75-99%	167	77.84%	75-99%	77	76.62%
	100%	151	83.44%	100%	179	83.24%
		891	33.00%		729	29.63%
Over-confident Respondents	26-49%	66	3.03%	26-49%	38	0.00%
	50%	431	3.71%	50%	173	0.58%
	51-74%	126	12.70%	51-74%	52	3.85%
	75-99%	622	28.14%	75-99%	220	33.18%
	100%	502	51.00%	100%	279	59.86%
		1747	26.62%		762	31.89%
Extremely Over-confident Respondents	75-99%	187	6.42%	75-99%	120	11.67%
	100%	565	15.04%	100%	270	15.19%
		752	12.90%		390	14.10%

Notes: Per the construction of the measure, the more confident individuals could not report below certain thresholds. This is the reason that 0 reported less than a 26% chance of completing college for the over-confident respondents. The same is true for under 75% for the extremely over-confident.

Chapter 4 Appendix

Table A.105: NLSY97 Income Summary Statistics

Variable	Full Sample	1997 Sub-Sample	2001 Sub-Sample
Income in 2012	\$ 38,574.00	\$ 42,358.20	\$ 38,697.80
Hours Worked in 2012	1794.65	1821.22	1813.99
Avg. Hourly Wage	\$23.26	\$24.92	\$23.09
Log Avg. Hrly Wage	2.86	2.93	2.84
College Expectation 1997		75.04	n/a
College Expectation 2001		n/a	70.57
Representative Sample	0.766	0.759	0.762
Age in 2012	29.87	31.35	29.83
Male	0.513	0.517	0.509
White	0.537	0.538	0.525
Black	0.253	0.25	0.25
HH Size	3.267	3.375	3.227
Married	0.454	0.518	0.439
ASVAB Percentile	48.314	48.227	48.96
–Imputed ASVAB	0.174	0.184	0.165
Years of Schooling	13.987	14.008	14.017
Potential Years of Experience	11.003	12.494	10.927
High School Degree	0.414	0.427	0.42
2 Year College Degree	0.084	0.085	0.078
4 Year College Degree	0.234	0.228	0.243
Masters Degree	0.077	0.085	0.073
Doctoral Degree	0.019	0.017	0.023
Rural	0.164	0.174	0.177
In MSA, Not Central City	0.526	0.537	0.507
In Central City MSA	0.425	0.408	0.432
North East Census	0.156	0.157	0.155
North Central Census	0.223	0.232	0.221
West Census	0.224	0.22	0.235
Agg, For, Fish Industry	0.007	0.005	0.007
Mining Industry	0.006	0.005	0.007
Utilities Industry	0.007	0.008	0.008
Construction Industry	0.058	0.066	0.053
Manufacturing Industry	0.068	0.073	0.053
Wholesale Trade Industry	0.025	0.03	0.017
Retail Trade Industry	0.098	0.093	0.098
Transportation and Warehouse Industry	0.035	0.034	0.038
Communications Industry	0.022	0.022	0.026
Finance, Insurance, and Real Estate Industry	0.065	0.07	0.058
Professional Industry	0.119	0.116	0.117
Education and Health Industry	0.226	0.233	0.235
Entertainment and Hospitality Industry	0.092	0.074	0.095
Other Service Industry	0.043	0.041	0.049
Public Administration Industry	0.041	0.047	0.044
Special ACS Industry	0.085	0.082	0.09
N	5205	1,887	1,143

Notes: Only those who reported positive income are utilized in the estimation and presented here. Average hourly wage was calculated as $\frac{\text{Yearly Income}}{\text{Hours Worked}}$.

Figure A.10: Frequency Distribution of College Completion Expectations

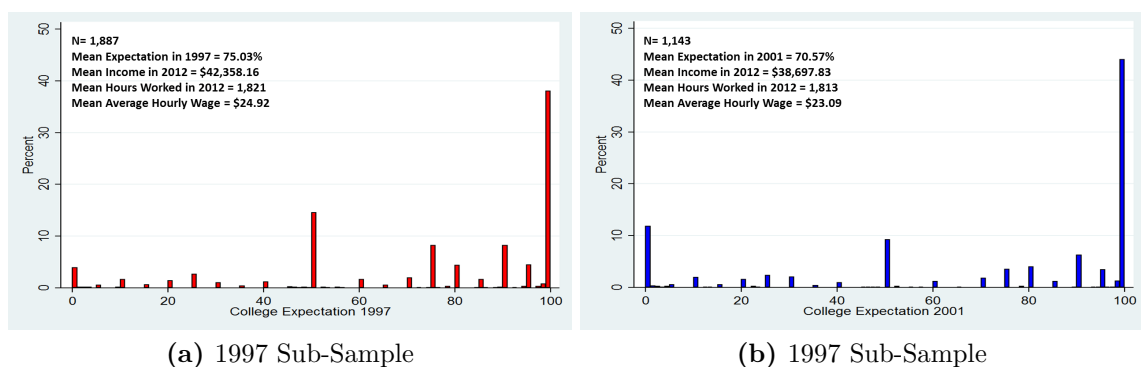


Table A.106: Income, Work, and Prediction Summary Statistics for 1997 Sub-sample

Variable	Full Sample	1997 Sub-sample	Successfully Predict College Completion	Unsuccessfully Predict College Completion	Successfully Predict No College Completion
Report Any Income in 2012	81.9% (38.5)	80.8% (39.4)	92.8% (25.9)	77.4% (41.8)	71.8% (45.1)
N	7092	2732	677	1080	436
Yearly Income in 2012	\$38,127.50 (30343.90)	\$41,745.80 (34114.70)	\$56,710.10 (39928.20)	\$36,763.00 (31321.70)	\$32,857.00 (27426.40)
N	5205	1997	583	744	274
Hours Worked in 2012	1622.32 (804.24)	1631.23 (819.67)	1868.09 (760.84)	1573.79 (808.21)	1452.46 (878.61)
N	6140	2374	634	917	360
Hourly Wage in 2012	23.18 (26.52)	24.87 (28.00)	31.83 (24.63)	21.16 (16.85)	20.02 (20.79)
N	4927	1891	565	700	256

Notes: Standard deviation in parenthesis. Those who reported between a 75% and 100% chance of completing a four-year college degree by the time they are 30 are classified as predicting college success, while those who reported lower than 50% are classified as not predicting college success.

Table A.107: Income, Work, and Prediction Summary Statistics for 2001 Sub-sample

Variable	Full Sample	2001 Sub-sample	Successfully Predict College Completion	Unsuccessfully Predict College Completion	Successfully Predict No College Completion
Report Any Income in 2012	81.9% (38.5)	82% (38.2)	94% (23.6)	82% (38.8)	73% (44.5)
N	7092	1630	439	531	436
Yearly Income in 2012	\$38,127.50 (30343.90)	\$38,362.10 (30845.3)	\$52,721.00 (35040.4)	\$33,174.60 (26344.9)	\$29,226.50 (26184.1)
N	5205	1211	386	391	277
Hours Worked in 2012	1622.32 (804.24)	1642.31 (816.53)	1835.66 (783.76)	1562.32 (772.55)	1537.84 (880.07)
N	6140	1415	414	456	352
Hourly Wage in 2012	23.18 (26.52)	22.991 (24.79)	30.834 (30.23)	20.464 (20.11)	17.166 (12.96)
N	4927	1148	373	362	264

Notes: Standard deviation in parenthesis. Those who reported between a 75% and 100% chance of completing a four-year college degree by the time they are 30 are classified as predicting college success, while those who reported lower than 50% are classified as not predicting college success.

Table A.108: Earnings Model Description

<u>Model Estimated:</u>	$Ln(Avg.HourlyWage_{i,2012}) = \alpha + \beta Expectation_{i,t} + \theta X_{i,2012} + \epsilon_i$	
<u>Variable of Interest:</u>	The percent chance the student believes that they will obtain a 4 year college degree by the time they are 30	
<u>Control Variables:</u>	Demographics	Location
	- Male	- Urban
	- White	- MSA Central City
	- Black	- MSA Non-Central City
		- 3 Census Tracks (W, NE, NC)
	Ability, Experience, and Household	
	- ASVAB Percentile	
	- Experience and Experience ²	
	- Marital Status	
	Schooling	Industry
	- Years of Schooling	- 1 of the 16 primary ACS industry codes for the respondent's primary employer
	- Indicators for Degree Attainment	

Notes: In the model, the t indicates which sub-sample that is used when estimating. It is not meant to indicate a panel analysis. For each categorical variable, that which is not listed is used as the base case.

Table A.109: Hourly Wage and Expectations Estimates

$Y = \ln(Avg.HourlyWage_{i,2012})$							
	<u>1997 Sub-sample</u>			<u>2001 Sub-sample</u>		<u>Overlapping Sub-sample</u>	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
College Expectation 1997			0.000902 (0.000606)				-0.000856 (0.00123)
College Expectation 2001					0.00203** (0.000682)		0.00220* (0.000995)
Male	0.150*** (0.0206)	0.160*** (0.0330)	0.164*** (0.0330)	0.191*** (0.0453)	0.197*** (0.0452)	0.202** (0.0758)	0.208** (0.0756)
White	0.000370 (0.0311)	0.0530 (0.0488)	0.0559 (0.0489)	0.00361 (0.0647)	0.0143 (0.0648)	0.131 (0.0974)	0.147 (0.0984)
Black	-0.114*** (0.0308)	-0.0904 (0.0478)	-0.0943* (0.0478)	-0.0354 (0.0652)	-0.0384 (0.0649)	-0.108 (0.0997)	-0.109 (0.100)
Experience	0.0405** (0.0136)	0.0677 (0.0367)	0.0707 (0.0367)	0.0362 (0.0339)	0.0448 (0.0338)	0.00786 (0.106)	0.0223 (0.105)
Experience ²	0.0000332 (0.000606)	-0.000672 (0.00116)	-0.000717 (0.00116)	0.000207 (0.00158)	0.00000906 (0.00156)	-0.00121 (0.00342)	-0.00168 (0.00338)
Married	0.160*** (0.0204)	0.167*** (0.0320)	0.164*** (0.0322)	0.229*** (0.0452)	0.236*** (0.0450)	0.177* (0.0759)	0.194* (0.0774)
ASVAB Percentile	0.00288*** (0.000469)	0.00313*** (0.000781)	0.00303*** (0.000784)	0.00321** (0.00100)	0.00286** (0.00101)	0.000761 (0.00195)	0.000205 (0.00200)
Years of Schooling	0.0504*** (0.00918)	0.0549* (0.0274)	0.0549* (0.0274)	0.0551** (0.0203)	0.0513* (0.0201)	-0.00210 (0.0610)	-0.00803 (0.0606)
High School Degree	0.148*** (0.0335)	0.174** (0.0535)	0.168** (0.0538)	0.168* (0.0792)	0.152 (0.0787)	0.166 (0.128)	0.176 (0.128)
2 Year College Degree	0.259*** (0.0487)	0.320*** (0.0807)	0.313*** (0.0809)	0.310** (0.109)	0.285** (0.108)	0.194 (0.173)	0.184 (0.171)
4 Year College Degree	0.391*** (0.0499)	0.414*** (0.0833)	0.401*** (0.0836)	0.349** (0.115)	0.301** (0.114)	0.309 (0.216)	0.285 (0.215)
Masters Degree	0.475*** (0.0658)	0.486*** (0.111)	0.476*** (0.111)	0.462** (0.148)	0.434** (0.146)	0.514 (0.289)	0.516 (0.286)
Doctoral Degree	0.629*** (0.0951)	0.882*** (0.162)	0.877*** (0.163)	0.626*** (0.187)	0.602** (0.186)	0.824* (0.332)	0.850* (0.331)
Industry Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4853	1860	1860	1129	1129	415	415
Adjusted R^2	0.218	0.214	0.215	0.231	0.237	0.243	0.246

Notes: Industry controls include which 1 of the 16 primary ACS industry codes the respondent's primary employer was classified as in 2012. Location controls include: whether the respondent lived in a rural or urban area; whether or not they live in an MSA, and if so if it were in the central city or not; and which of the 4 primary census regions the respondents lived in at the time of the survey.

- Robust Robust standard errors in parentheses

- * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.110: Hourly Wage and Expectations - Men and Women

$Y = \ln(Avg.HourlyWage_{i,2012})$						
	<u>Full Sample</u>		<u>1997 Sub-sample</u>		<u>2001 Sub-sample</u>	
	(1-M)	(1-F)	(2-M)	(2-F)	(3-M)	(3-F)
College Expectation 1997			0.00101 (0.000915)	0.000799 (0.000878)		
College Expectation 2001					0.00214* (0.000950)	0.00224* (0.00103)
White	-0.00979 (0.0440)	0.0108 (0.0488)	0.0243 (0.0685)	0.102 (0.0764)	-0.0103 (0.0905)	0.00524 (0.100)
Black	-0.201*** (0.0485)	-0.0545 (0.0441)	-0.193* (0.0784)	-0.0174 (0.0666)	-0.188 (0.108)	0.0601 (0.0811)
Experience	0.0344 (0.0203)	0.0454* (0.0201)	0.0654 (0.0592)	0.0840 (0.0568)	0.112* (0.0455)	-0.0277 (0.0399)
Experience ²	0.000189 (0.000864)	-0.000111 (0.000946)	0.000747 (0.00164)	-0.00242 (0.00198)	-0.00274 (0.00194)	0.00320 (0.00170)
Married	0.211*** (0.0311)	0.111*** (0.0280)	0.216*** (0.0500)	0.115** (0.0443)	0.325*** (0.0717)	0.119 (0.0610)
ASVAB Percentile	0.00179* (0.000699)	0.00438*** (0.000702)	0.00219 (0.00118)	0.00402*** (0.00120)	0.00113 (0.00160)	0.00455** (0.00143)
Years of Schooling	0.0531*** (0.0134)	0.0482*** (0.0135)	0.105* (0.0452)	0.0114 (0.0367)	0.0718* (0.0320)	0.0442 (0.0260)
High School Degree	0.191*** (0.0469)	0.0921 (0.0516)	0.168* (0.0821)	0.128 (0.0719)	0.171 (0.106)	0.0601 (0.117)
2 Year College Degree	0.197** (0.0664)	0.311*** (0.0761)	0.232 (0.125)	0.378*** (0.111)	0.331* (0.159)	0.189 (0.149)
4 Year College Degree	0.365*** (0.0699)	0.404*** (0.0745)	0.299* (0.125)	0.471*** (0.119)	0.329 (0.171)	0.166 (0.160)
Masters Degree	0.309* (0.123)	0.511*** (0.0922)	0.0504 (0.231)	0.672*** (0.150)	0.482* (0.221)	0.229 (0.201)
Doctoral Degree	0.557*** (0.143)	0.655*** (0.130)	0.555* (0.271)	1.053*** (0.200)	0.657* (0.284)	0.448 (0.246)
Industry Controls	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2483	2370	959	901	576	553
Adjusted R^2	0.169	0.244	0.154	0.241	0.221	0.244

Notes: Industry controls include which 1 of the 16 primary ACS industry codes the respondent's primary employer was classified as in 2012 . Location controls include: whether the respondent lived in a rural or urban area; whether or not they live in an MSA, and if so if it were in the central city or not; and which of the 4 primary census regions the respondents lived in at the time of the survey.

Robust standard errors in parentheses

– * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.111: Quantile and OLS Estimates of Hourly Wage and Expectations

$Y = \ln(\text{Avg.HourlyWage}_{i,2012})$								
	1997 Sub-sample				2001 Sub-sample			
	OLS	$\tau = .25$	$\tau = .5$ (Median)	$\tau = .75$	OLS	$\tau = .25$	$\tau = .5$ (Median)	$\tau = .75$
College Expectation 1997	0.000912 (0.000623)	0.000641 (0.000470)	0.000486 (0.000432)	0.00148** (0.000541)				
College Expectation 2001					0.00201** (0.000691)	0.00156** (0.000533)	0.00191*** (0.000554)	0.00188*** (0.000545)
Male	0.150*** (0.0360)	0.118*** (0.0281)	0.171*** (0.0286)	0.206*** (0.0347)	0.191*** (0.0460)	0.145*** (0.0319)	0.184*** (0.0346)	0.251*** (0.0345)
White	0.0598 (0.0509)	0.0329 (0.0409)	0.0263 (0.0381)	0.0449 (0.0478)	0.0152 (0.0658)	-0.0490 (0.0583)	-0.0238 (0.0456)	0.0677 (0.0527)
Black	-0.113* (0.0522)	-0.0368 (0.0362)	-0.0622 (0.0390)	0.00825 (0.0450)	-0.0474 (0.0679)	-0.133* (0.0579)	-0.0117 (0.0450)	0.116* (0.0516)
Experience	0.0772 (0.0395)	0.0607 (0.0329)	-0.00444 (0.0336)	0.0259 (0.0294)	0.0416 (0.0341)	0.0700 (0.0426)	0.0433 (0.0288)	0.0503* (0.0208)
Experience ²	-0.000936 (0.00124)	-0.000263 (0.00118)	0.000387 (0.00108)	0.000424 (0.000734)	0.0000628 (0.00155)	-0.00142 (0.00213)	0.000273 (0.00144)	-0.0000356 (0.00101)
Married	0.162*** (0.0329)	0.153*** (0.0282)	0.137*** (0.0275)	0.133*** (0.0333)	0.233*** (0.0466)	0.274*** (0.0326)	0.186*** (0.0354)	0.179*** (0.0393)
ASVAB Percentile	0.00288*** (0.000839)	0.00295*** (0.000637)	0.00323*** (0.000656)	0.00425*** (0.000775)	0.00279* (0.00109)	0.00225** (0.000802)	0.00281*** (0.000773)	0.00375*** (0.000848)
Years of Schooling	0.0556 (0.0291)	0.0617** (0.0226)	0.00830 (0.0226)	0.0371 (0.0270)	0.0497* (0.0206)	0.0384* (0.0185)	0.0533** (0.0163)	0.0275 (0.0173)
High School Degree	0.167** (0.0549)	0.188*** (0.0480)	0.160*** (0.0397)	0.137** (0.0493)	0.158 (0.0814)	0.159* (0.0733)	0.0921 (0.0634)	0.150* (0.0738)
2 Year College Degree	0.313*** (0.0828)	0.278*** (0.0736)	0.312*** (0.0641)	0.385*** (0.0704)	0.290** (0.112)	0.339** (0.105)	0.192* (0.0830)	0.354*** (0.0969)
4 Year College Degree	0.409*** (0.0849)	0.400*** (0.0723)	0.418*** (0.0727)	0.377*** (0.0775)	0.307** (0.117)	0.374*** (0.108)	0.280** (0.0877)	0.459*** (0.0962)
Masters Degree	0.445*** (0.121)	0.527*** (0.0893)	0.439*** (0.0978)	0.441*** (0.0991)	0.440** (0.147)	0.510*** (0.116)	0.366*** (0.110)	0.638*** (0.114)
Doctoral Degree	0.892*** (0.166)	0.803*** (0.158)	0.952*** (0.194)	0.956*** (0.141)	0.610** (0.187)	0.594* (0.301)	0.579*** (0.147)	0.731** (0.226)
Industry Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1860	1860	1860	1860	1129	1129	1129	1129
Adjusted R^2	0.193				0.221			
Pseudo R^2		0.1461	0.1437	0.1379		0.1719	0.1713	0.1595

Notes: Industry controls include which 1 of the 16 primary ACS industry codes the respondent's primary employer was classified as in 2012. Location controls include: whether the respondent lived in a rural or urban area; whether or not they live in an MSA, and if so if it were in the central city or not; and which of the 4 primary census regions the respondents lived in at the time of the survey.

Robust standard errors in parentheses

– * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.112: Quantile and OLS Estimates of Average Hourly Wage and Expectations
- 1997 Sub-sample

$Y = \ln(Avg.HourlyWage_{i,2012})$								
	Male Respondents				Female Respondents			
	OLS	$\tau = .25$	$\tau = .5$ (Median)	$\tau = .75$	OLS	$\tau = .25$	$\tau = .5$ (Median)	$\tau = .75$
College Expectation 1997	0.00101 (0.000915)	0.00106 (0.000757)	-0.000166 (0.000612)	-0.00132 (0.000877)	0.000799 (0.000878)	-0.000501 (0.000656)	0.00111* (0.000548)	0.00198** (0.000752)
White	0.0243 (0.0685)	0.00643 (0.0675)	-0.0281 (0.0568)	-0.0124 (0.0605)	0.102 (0.0764)	-0.0156 (0.0593)	0.0632 (0.0551)	0.146** (0.0552)
Black	-0.193* (0.0784)	-0.259*** (0.0690)	-0.0663 (0.0597)	0.00950 (0.0704)	-0.0174 (0.0666)	0.0523 (0.0510)	0.00649 (0.0529)	0.0267 (0.0504)
Experience	0.0654 (0.0592)	0.0880 (0.0521)	-0.0123 (0.0402)	-0.0252 (0.0529)	0.0840 (0.0568)	0.0702 (0.0425)	0.0202 (0.0504)	0.00127 (0.0856)
Experience ²	0.000747 (0.00164)	0.000710 (0.00157)	0.00145 (0.00117)	0.00161 (0.00139)	-0.00242 (0.00198)	-0.00138 (0.00160)	-0.00175* (0.000878)	0.000539 (0.00152)
Married	0.216*** (0.0500)	0.272*** (0.0443)	0.164*** (0.0383)	0.129* (0.0541)	0.115** (0.0443)	0.0921* (0.0425)	0.135*** (0.0370)	0.0836* (0.0368)
ASVAB Percentile	0.00219 (0.00118)	0.00173 (0.000940)	0.00265** (0.000886)	0.00297* (0.00126)	0.00402*** (0.00120)	0.00526*** (0.000787)	0.00354*** (0.000845)	0.00386*** (0.000799)
Years of Schooling	0.105* (0.0452)	0.125** (0.0385)	0.0415 (0.0308)	0.0359 (0.0406)	0.0114 (0.0367)	0.0261 (0.0258)	-0.0325 (0.0479)	0.00861 (0.0766)
High School Degree	0.168* (0.0821)	0.177** (0.0552)	0.239*** (0.0565)	0.184** (0.0651)	0.128 (0.0719)	0.193 (0.107)	0.0755 (0.0669)	0.103 (0.0722)
2 Year College Degree	0.232 (0.125)	0.138 (0.0967)	0.302** (0.101)	0.333** (0.122)	0.378*** (0.111)	0.486*** (0.122)	0.357*** (0.0833)	0.369*** (0.0941)
4 Year College Degree	0.299* (0.125)	0.285** (0.0978)	0.406*** (0.0931)	0.302*** (0.0858)	0.471*** (0.119)	0.519*** (0.134)	0.417*** (0.0891)	0.388*** (0.0951)
Masters Degree	0.0504 (0.231)	0.361* (0.158)	0.400** (0.131)	0.211 (0.163)	0.672*** (0.150)	0.689*** (0.149)	0.509*** (0.104)	0.471*** (0.137)
Doctoral Degree	0.555* (0.271)	0.677 (0.456)	0.574 (0.321)	0.932* (0.419)	1.053*** (0.200)	0.941*** (0.242)	1.024*** (0.116)	0.924*** (0.137)
Industry Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	959	959	959	959	901	901	901	901
Adjusted R^2	0.154				0.241			
Pseudo R^2		0.137	0.131	0.111		0.179	0.179	0.183

Notes: Industry controls include which 1 of the 16 primary ACS industry codes the respondent's primary employer was classified as in 2012. Location controls include: whether the respondent lived in a rural or urban area; whether or not they live in an MSA, and if so if it were in the central city or not; and which of the 4 primary census regions the respondents lived in at the time of the survey.

Robust standard errors in parentheses
- * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.113: Quantile Estimates of Average Hourly Wage and Expectations - 2001 Sub-sample

$Y = \text{Ln}(\text{Avg.HourlyWage}_{i,2012})$								
	Male Respondents				Female Respondents			
	OLS	$\tau = .25$	$\tau = .5$ (Median)	$\tau = .75$	OLS	$\tau = .25$	$\tau = .5$ (Median)	$\tau = .75$
College Expectation 2001	0.00214* (0.000950)	0.00126 (0.000786)	0.00146 (0.000775)	0.00262** (0.000934)	0.00224* (0.00103)	0.00245** (0.000851)	0.00116 (0.000669)	0.00183 (0.00122)
White	-0.0103 (0.0905)	-0.0801 (0.0825)	-0.0148 (0.0697)	0.0232 (0.0820)	0.00524 (0.100)	-0.0932 (0.0754)	-0.0585 (0.0588)	0.103 (0.0827)
Black	-0.188 (0.108)	-0.184* (0.0782)	-0.164* (0.0782)	0.0196 (0.0911)	0.0601 (0.0811)	-0.0430 (0.0965)	0.112** (0.0378)	0.115 (0.0816)
Experience	0.112* (0.0455)	0.143 (0.0807)	0.0881* (0.0427)	0.106* (0.0476)	-0.0277 (0.0399)	-0.00809 (0.0316)	0.000627 (0.0275)	0.0208 (0.0394)
Experience ²	-0.00274 (0.00194)	-0.00478 (0.00360)	-0.00174 (0.00200)	-0.00273 (0.00232)	0.00320 (0.00170)	0.00240 (0.00151)	0.00211 (0.00126)	0.00153 (0.00182)
Married	0.325*** (0.0717)	0.400*** (0.0679)	0.230*** (0.0529)	0.264*** (0.0553)	0.119 (0.0610)	0.0546 (0.0462)	0.167*** (0.0418)	0.0935 (0.0627)
ASVAB Percentile	0.00113 (0.00160)	0.00154 (0.00122)	0.000516 (0.00120)	0.000557 (0.00136)	0.00455** (0.00143)	0.00456*** (0.00120)	0.00493*** (0.00105)	0.00511*** (0.00140)
Years of Schooling	0.0718* (0.0320)	0.0652 (0.0412)	0.0618* (0.0272)	0.0602* (0.0285)	0.0442 (0.0260)	0.0187 (0.0244)	0.0413* (0.0171)	0.0387 (0.0315)
High School Degree	0.171 (0.106)	0.0534 (0.123)	0.175* (0.0876)	0.0271 (0.0812)	0.0601 (0.117)	0.207* (0.102)	0.0132 (0.0817)	0.104 (0.144)
2 Year College Degree	0.331* (0.159)	0.221 (0.188)	0.322* (0.140)	0.0679 (0.143)	0.189 (0.149)	0.478*** (0.126)	0.282* (0.118)	0.392* (0.175)
4 Year College Degree	0.329 (0.171)	0.235 (0.185)	0.389** (0.149)	0.179 (0.139)	0.166 (0.160)	0.472** (0.146)	0.258* (0.122)	0.398* (0.198)
Masters Degree	0.482* (0.221)	0.419 (0.229)	0.465* (0.207)	0.369 (0.205)	0.229 (0.201)	0.581** (0.188)	0.281 (0.161)	0.541* (0.236)
Doctoral Degree	0.657* (0.284)	0.444 (0.295)	0.691** (0.223)	0.807* (0.407)	0.448 (0.246)	0.659*** (0.197)	0.519* (0.221)	0.666 (0.456)
Industry Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	576	576	576	576	553	553	553	553
Adjusted R^2	0.221				0.244			
Pseudo R^2		0.2033	0.1766	0.1525		0.2030	0.1960	0.1929

Notes: Industry controls include which 1 of the 16 primary ACS industry codes the respondent's primary employer was classified as in 2012 . Location controls include: whether the respondent lived in a rural or urban area; whether or not they live in an MSA, and if so if it were in the central city or not; and which of the 4 primary census regions the respondents lived in at the time of the survey.

Robust standard errors in parentheses

– * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.114: Average Hourly Wage and Expectations - Expectations with Indicators

$Y = Ln(Avg.HourlyWage_{i,2012})$								
	<u>1997 Sub-sample</u>				<u>2001 Sub-sample</u>			
	OLS	$\tau = .25$	$\tau = .5$ (Median)	$\tau = .75$	OLS	$\tau = .25$	$\tau = .5$ (Median)	$\tau = .75$
College Expectation 1997	0.000239 (0.000868)	-0.000833 (0.000622)	-0.000450 (0.000642)	0.000964 (0.000635)				
Expectation =0	-0.00715 (0.116)	-0.0727 (0.0554)	-0.106 (0.0719)	0.255 (0.196)				
Expectation = 50	0.00252 (0.0513)	-0.0392 (0.0373)	0.0151 (0.0379)	0.00441 (0.0326)				
Expectation = 100	0.0661 (0.0434)	0.112** (0.0347)	0.0517 (0.0364)	0.0572 (0.0404)				
College Expectation 2001					0.00165 (0.00119)	0.00197* (0.000909)	0.000482 (0.000922)	0.00126 (0.000914)
Expectation =0					0.0601 (0.105)	0.101 (0.0810)	-0.0630 (0.0868)	-0.0111 (0.0894)
Expectation = 50					-0.0232 (0.109)	0.0622 (0.0786)	0.00690 (0.0489)	-0.0487 (0.0827)
Expectation = 100					0.0941 (0.0591)	0.0414 (0.0486)	0.0895* (0.0410)	0.0887* (0.0431)
Male	0.155*** (0.0359)	0.137*** (0.0282)	0.170*** (0.0276)	0.209*** (0.0324)	0.196*** (0.0460)	0.151*** (0.0389)	0.178*** (0.0334)	0.259*** (0.0369)
White	0.0607 (0.0509)	0.0261 (0.0347)	0.0338 (0.0370)	0.0167 (0.0463)	0.0167 (0.0661)	-0.0183 (0.0534)	-0.00244 (0.0440)	0.0714 (0.0599)
Black	-0.116* (0.0520)	-0.0428 (0.0411)	-0.0547 (0.0387)	-0.0111 (0.0413)	-0.0496 (0.0684)	-0.125* (0.0566)	0.00673 (0.0370)	0.107 (0.0581)
Experience	0.0768 (0.0394)	0.0762* (0.0311)	-0.000748 (0.0303)	0.0242 (0.0299)	0.0390 (0.0343)	0.0714 (0.0418)	0.0329 (0.0300)	0.0459* (0.0226)
Experience ²	-0.000916 (0.00124)	-0.000731 (0.00102)	0.000194 (0.000964)	0.000679 (0.000872)	-0.00000410 (0.00156)	-0.00166 (0.00209)	0.000639 (0.00150)	-0.000116 (0.00106)
Married	0.160*** (0.0330)	0.151*** (0.0254)	0.129*** (0.0274)	0.130*** (0.0301)	0.234*** (0.0462)	0.250*** (0.0390)	0.190*** (0.0349)	0.187*** (0.0395)
ASVAB Percentile	0.00295*** (0.000841)	0.00333*** (0.000601)	0.00341*** (0.000630)	0.00418*** (0.000727)	0.00279* (0.00109)	0.00228** (0.000798)	0.00270*** (0.000779)	0.00338*** (0.000801)
Years of Schooling	0.0560 (0.0291)	0.0647** (0.0217)	0.00832 (0.0214)	0.0453 (0.0246)	0.0446* (0.0209)	0.0343 (0.0209)	0.0554*** (0.0157)	0.0195 (0.0173)
High School Degree	0.168** (0.0555)	0.186*** (0.0519)	0.149*** (0.0409)	0.134** (0.0424)	0.150 (0.0830)	0.197* (0.0770)	0.0791 (0.0626)	0.136 (0.0734)
2 Year College Degree	0.315*** (0.0836)	0.288*** (0.0778)	0.296*** (0.0620)	0.362*** (0.0680)	0.283* (0.112)	0.384*** (0.109)	0.161* (0.0783)	0.347*** (0.101)
4 Year College Degree	0.406*** (0.0850)	0.409*** (0.0755)	0.398*** (0.0681)	0.366*** (0.0701)	0.291* (0.116)	0.413*** (0.114)	0.246** (0.0877)	0.446*** (0.0976)
Masters Degree	0.436*** (0.121)	0.519*** (0.0971)	0.428*** (0.0884)	0.390*** (0.104)	0.418** (0.148)	0.552*** (0.121)	0.338** (0.118)	0.636*** (0.119)
Doctoral Degree	0.888*** (0.167)	0.762*** (0.134)	0.956*** (0.190)	0.908*** (0.152)	0.598** (0.186)	0.709* (0.322)	0.530*** (0.140)	0.739** (0.254)
Industry Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1860	1860	1860	1860	1129	1129	1129	1129
Adjusted R^2	0.193				0.221			
Pseudo R^2		0.149	0.145	.140		0.173	0.173	0.161

Notes: Industry controls include which 1 of the 16 primary ACS industry codes the respondent's primary employer was classified as in 2012 . Location controls include: whether the respondent lived in a rural or urban area; whether or not they live in an MSA, and if so if it were in the central city or not; and which of the 4 primary census regions the respondents lived in at the time of the survey.

Robust standard errors in parentheses

– * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.115: Average Hourly Wage and Expectations - 1997 Expectations with Indicators

$Y = \ln(\text{Avg.HourlyWage}_{i,2012})$								
	Male Respondents				Female Respondents			
	OLS	$\tau = .25$	$\tau = .5$ (Median)	$\tau = .75$	OLS	$\tau = .25$	$\tau = .5$ (Median)	$\tau = .75$
College Expectation 1997	0.000246 (0.00120)	-0.00189* (0.000897)	-0.00177* (0.000785)	-0.000136 (0.00116)	0.000243 (0.00126)	-0.000773 (0.000966)	0.000997 (0.000848)	0.00146 (0.00115)
Expectation =0	0.119 (0.153)	-0.117 (0.0994)	-0.0460 (0.165)	0.331*** (0.0965)	-0.230 (0.180)	-0.134 (0.120)	-0.134 (0.102)	-0.286 (0.232)
Expectation = 50	0.0299 (0.0706)	-0.00456 (0.0649)	0.0311 (0.0467)	0.0549 (0.0573)	-0.0326 (0.0698)	-0.0426 (0.0573)	0.00636 (0.0464)	-0.0827 (0.0564)
Expectation = 100	0.154* (0.0693)	0.249*** (0.0401)	0.145** (0.0450)	0.129 (0.0889)	-0.0236 (0.0551)	-0.0293 (0.0479)	-0.0323 (0.0480)	-0.0244 (0.0388)
White	0.0289 (0.0682)	0.0477 (0.0500)	0.00491 (0.0542)	-0.0655 (0.0787)	0.107 (0.0762)	-0.00550 (0.0595)	0.0875 (0.0512)	0.138** (0.0475)
Black	-0.209** (0.0779)	-0.214*** (0.0634)	-0.0835 (0.0545)	-0.0127 (0.0844)	-0.0135 (0.0669)	0.0661 (0.0559)	-0.00948 (0.0516)	0.00682 (0.0466)
Experience	0.0655 (0.0590)	0.0704 (0.0364)	-0.00252 (0.0434)	0.0108 (0.0742)	0.0839 (0.0567)	0.0751 (0.0836)	0.0361 (0.0351)	0.0240 (0.0355)
Experience ²	0.000805 (0.00164)	0.000537 (0.00122)	0.00181 (0.00129)	0.00111 (0.00245)	-0.00232 (0.00198)	-0.00175 (0.00122)	-0.00147 (0.000808)	0.0000634 (0.00129)
Married	0.213*** (0.0502)	0.249*** (0.0365)	0.165*** (0.0390)	0.165** (0.0518)	0.114* (0.0443)	0.0994* (0.0416)	0.145*** (0.0395)	0.0497 (0.0353)
ASVAB Percentile	0.00231 (0.00118)	0.00195* (0.000776)	0.00310*** (0.000903)	0.00277* (0.00115)	0.00400*** (0.00120)	0.00564*** (0.000748)	0.00303*** (0.000917)	0.00382*** (0.000744)
Years of Schooling	0.107* (0.0448)	0.106*** (0.0300)	0.0668* (0.0295)	0.0462 (0.0449)	0.0134 (0.0369)	0.0187 (0.0764)	-0.00460 (0.0318)	0.0191 (0.0282)
High School Degree	0.174* (0.0831)	0.188** (0.0633)	0.227*** (0.0574)	0.226*** (0.0677)	0.118 (0.0726)	0.158 (0.104)	0.0574 (0.0648)	0.0786 (0.0675)
2 Year College Degree	0.248* (0.126)	0.146 (0.0869)	0.243* (0.0994)	0.401* (0.167)	0.371*** (0.111)	0.462*** (0.127)	0.328*** (0.0946)	0.357*** (0.0927)
4 Year College Degree	0.297* (0.125)	0.299** (0.0956)	0.346*** (0.0986)	0.372* (0.146)	0.469*** (0.119)	0.520*** (0.130)	0.380*** (0.0921)	0.379*** (0.0949)
Masters Degree	0.0320 (0.230)	0.363** (0.127)	0.280* (0.131)	0.293 (0.232)	0.670*** (0.149)	0.675*** (0.148)	0.473*** (0.106)	0.506*** (0.134)
Doctoral Degree	0.538 (0.275)	0.638 (0.515)	0.546 (0.317)	1.046** (0.358)	1.042*** (0.201)	0.892*** (0.230)	0.991*** (0.130)	0.930*** (0.135)
Industry Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	959	959	959	959	901	901	901	901
Adjusted R^2	0.158				0.241			
Pseudo R^2		0.147	0.135	0.120		0.181	0.181	0.185

Notes: Industry controls include which 1 of the 16 primary ACS industry codes the respondent's primary employer was classified as in 2012. Location controls include: whether the respondent lived in a rural or urban area; whether or not they live in an MSA, and if so if it were in the central city or not; and which of the 4 primary census regions the respondents lived in at the time of the survey.

Robust standard errors in parentheses

- * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.116: Average Hourly Wage and Expectations - 2001 Expectations with Indicators

$Y = \ln(Avg.HourlyWage_{i,2012})$								
	Male Respondents				Female Respondents			
	OLS	$\tau = .25$	$\tau = .5$ (Median)	$\tau = .75$	OLS	$\tau = .25$	$\tau = .5$ (Median)	$\tau = .75$
College Expectation 2001	0.00171 (0.00144)	0.000694 (0.00186)	0.00108 (0.000943)	-0.000207 (0.00133)	0.00270 (0.00194)	0.00222 (0.00155)	0.000699 (0.00117)	0.00183 (0.00157)
Expectation =0	-0.0133 (0.130)	-0.0632 (0.156)	-0.126 (0.0807)	-0.210* (0.101)	0.208 (0.172)	0.0355 (0.150)	0.0622 (0.115)	0.264 (0.185)
Expectation = 50	-0.131 (0.158)	-0.0132 (0.116)	-0.0938 (0.112)	-0.0145 (0.0965)	0.160 (0.124)	0.155 (0.0989)	0.128 (0.0806)	0.0312 (0.0911)
Expectation = 100	0.0496 (0.0854)	0.0220 (0.0804)	-0.0418 (0.0606)	0.177* (0.0697)	0.127 (0.0838)	0.147** (0.0459)	0.189*** (0.0486)	0.0750 (0.0664)
White	-0.00794 (0.0918)	-0.0737 (0.0980)	-0.0190 (0.0598)	0.0462 (0.0865)	0.00184 (0.0995)	-0.0592 (0.0527)	-0.0393 (0.0525)	0.0872 (0.0760)
Black	-0.200 (0.111)	-0.190* (0.0912)	-0.196** (0.0730)	0.0129 (0.0942)	0.0558 (0.0812)	-0.0163 (0.0581)	0.108* (0.0470)	0.120 (0.0734)
Experience	0.114* (0.0452)	0.151* (0.0739)	0.100** (0.0378)	0.116* (0.0559)	-0.0323 (0.0402)	-0.0110 (0.0235)	0.0139 (0.0229)	0.0211 (0.0366)
Experience ²	-0.00285 (0.00197)	-0.00520 (0.00378)	-0.00169 (0.00168)	-0.00390 (0.00271)	0.00313 (0.00171)	0.00218* (0.000991)	0.00144 (0.00106)	0.00144 (0.00163)
Married	0.331*** (0.0699)	0.400*** (0.0686)	0.234*** (0.0514)	0.275*** (0.0587)	0.117 (0.0612)	0.101** (0.0367)	0.149*** (0.0372)	0.129* (0.0540)
ASVAB Percentile	0.000985 (0.00164)	0.00118 (0.00125)	0.000509 (0.00106)	0.00118 (0.00125)	0.00452** (0.00142)	0.00347*** (0.000929)	0.00455*** (0.000824)	0.00422*** (0.00124)
Years of Schooling	0.0695* (0.0335)	0.0651 (0.0349)	0.0728** (0.0256)	0.0373 (0.0264)	0.0383 (0.0257)	0.0147 (0.0189)	0.0348* (0.0172)	0.0396 (0.0235)
High School Degree	0.161 (0.105)	0.0498 (0.126)	0.149 (0.0899)	0.0747 (0.0837)	0.0702 (0.122)	0.152 (0.115)	0.0510 (0.0778)	0.0909 (0.106)
2 Year College Degree	0.325* (0.159)	0.212 (0.195)	0.285* (0.131)	0.116 (0.145)	0.198 (0.148)	0.370** (0.134)	0.284** (0.108)	0.405** (0.127)
4 Year College Degree	0.313 (0.168)	0.232 (0.194)	0.399** (0.148)	0.301* (0.119)	0.162 (0.163)	0.444** (0.144)	0.306** (0.108)	0.376* (0.161)
Masters Degree	0.468* (0.220)	0.422 (0.243)	0.473* (0.206)	0.505** (0.188)	0.216 (0.204)	0.487** (0.152)	0.319* (0.138)	0.567** (0.208)
Doctoral Degree	0.651* (0.282)	0.452 (0.286)	0.661** (0.230)	0.947** (0.352)	0.449 (0.247)	0.622** (0.192)	0.727*** (0.188)	0.574* (0.274)
Industry Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	576	576	576	576	553	553	553	553
Adjusted R^2	0.219				0.248			
Pseudo R^2		0.203	0.178	0.157		0.208	0.206	0.197

Notes: Industry controls include which 1 of the 16 primary ACS industry codes the respondent's primary employer was classified as in 2012 . Location controls include: whether the respondent lived in a rural or urban area; whether or not they live in an MSA, and if so if it were in the central city or not; and which of the 4 primary census regions the respondents lived in at the time of the survey.

Robust standard errors in parentheses

- * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.117: Average Hourly Wage and Non-Linear Expectations in 1997

	$Y = \ln(Avg.HourlyWage_{i,2012})$					
	<u>Full 1997 Sub-sample</u>		<u>Men in 1997 Sub-sample</u>		<u>Women in 1997 Sub-sample</u>	
	(1)	(2)	(3)	(4)	(5)	(6)
College Expectation 1997	-0.000949 (0.00254)	0.00910 (0.00618)	-0.00566 (0.00369)	0.0172* (0.00860)	0.00580 (0.00354)	0.00298 (0.00893)
Expectations ²	0.0000160 (0.0000203)	-0.000228 (0.000136)	0.0000588 (0.0000301)	-0.000504** (0.000193)	-0.0000419 (0.0000276)	0.0000252 (0.000192)
Expectations ³		0.00000151 (0.000000831)		0.00000352** (0.00000120)		-0.000000410 (0.00000114)
ASVAB Percentile	0.00285*** (0.000841)	0.00294*** (0.000840)	0.00202 (0.00118)	0.00222 (0.00118)	0.00406*** (0.00120)	0.00403*** (0.00120)
Experience	0.0785 (0.0414)	0.0797 (0.0414)	0.0594 (0.0616)	0.0694 (0.0611)	0.0902 (0.0598)	0.0910 (0.0599)
Experience ²	-0.000936 (0.00124)	-0.000950 (0.00124)	0.000860 (0.00164)	0.000690 (0.00162)	-0.00235 (0.00198)	-0.00238 (0.00198)
Male	0.152*** (0.0359)	0.156*** (0.0357)				
White	0.0592 (0.0509)	0.0616 (0.0508)	0.0263 (0.0684)	0.0276 (0.0684)	0.107 (0.0762)	0.105 (0.0762)
Black	-0.115* (0.0520)	-0.115* (0.0520)	-0.201* (0.0783)	-0.210** (0.0781)	-0.0110 (0.0668)	-0.0119 (0.0671)
Married	0.162*** (0.0329)	0.161*** (0.0329)	0.214*** (0.0501)	0.218*** (0.0495)	0.112* (0.0443)	0.113* (0.0445)
Years of Schooling	0.0550 (0.0292)	0.0559 (0.0292)	0.104* (0.0453)	0.112* (0.0449)	0.0146 (0.0368)	0.0149 (0.0369)
High School Degree	0.169** (0.0553)	0.170** (0.0553)	0.175* (0.0830)	0.172* (0.0826)	0.122 (0.0718)	0.122 (0.0718)
2 Year College Degree	0.316*** (0.0836)	0.315*** (0.0832)	0.249 (0.127)	0.242 (0.125)	0.370*** (0.111)	0.371*** (0.111)
4 Year College Degree	0.408*** (0.0848)	0.408*** (0.0847)	0.293* (0.125)	0.287* (0.123)	0.469*** (0.119)	0.470*** (0.119)
Masters Degree	0.442*** (0.121)	0.437*** (0.121)	0.0284 (0.232)	-0.0104 (0.231)	0.670*** (0.150)	0.671*** (0.150)
Doctoral Degree	0.890*** (0.166)	0.889*** (0.166)	0.535 (0.274)	0.508 (0.276)	1.051*** (0.199)	1.050*** (0.199)
Industry Controls	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1860	1860	959	959	901	901
Adjusted R^2	0.193	0.194	0.158	0.165	0.243	0.242

Notes: Industry controls include which 1 of the 16 primary ACS industry codes the respondent's primary employer was classified as in 2012. Location controls include: whether the respondent lived in a rural or urban area; whether or not they live in an MSA, and if so if it were in the central city or not; and which of the 4 primary census regions the respondents lived in at the time of the survey.

Robust standard errors in parentheses

– * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.118: Average Hourly Wage and Non-Linear Expectations in 2001

	$Y = \ln(\text{Avg.HourlyWage}_{i,2012})$					
	<u>Full 2001 Sub-sample</u>		<u>Men in 2001 Sub-sample</u>		<u>Women in 2001 Sub-sample</u>	
	(1)	(2)	(3)	(4)	(5)	(6)
College Expectation 2001	-0.00151 (0.00317)	-0.00425 (0.00697)	-0.00166 (0.00474)	-0.00833 (0.00926)	-0.000463 (0.00422)	-0.000438 (0.0103)
Expectations ²	0.0000339 (0.0000294)	0.000108 (0.000168)	0.0000373 (0.0000451)	0.000219 (0.000223)	0.0000255 (0.0000380)	0.0000249 (0.000252)
Expectations ³		-0.000000479 (0.00000106)		-0.00000119 (0.00000143)		4.32e-09 (0.00000160)
ASVAB Percentile	0.00278* (0.00109)	0.00281** (0.00109)	0.00108 (0.00160)	0.00117 (0.00160)	0.00457** (0.00143)	0.00457** (0.00143)
Experience	0.0423 (0.0370)	0.0432 (0.0371)	0.120* (0.0486)	0.123* (0.0485)	-0.0339 (0.0431)	-0.0339 (0.0433)
Experience ²	-0.0000731 (0.00156)	-0.0000985 (0.00156)	-0.00296 (0.00196)	-0.00307 (0.00196)	0.00312 (0.00171)	0.00312 (0.00172)
Male	0.194*** (0.0458)	0.194*** (0.0458)				
White	0.0131 (0.0658)	0.0128 (0.0659)	-0.0124 (0.0910)	-0.0176 (0.0915)	0.00300 (0.0997)	0.00299 (0.0996)
Black	-0.0508 (0.0683)	-0.0505 (0.0684)	-0.195 (0.110)	-0.198 (0.110)	0.0598 (0.0811)	0.0598 (0.0817)
Married	0.232*** (0.0467)	0.233*** (0.0467)	0.327*** (0.0709)	0.330*** (0.0713)	0.117 (0.0612)	0.117 (0.0617)
Years of Schooling	0.0465* (0.0212)	0.0470* (0.0211)	0.0672* (0.0340)	0.0679* (0.0339)	0.0423 (0.0259)	0.0423 (0.0259)
High School Degree	0.149 (0.0823)	0.147 (0.0819)	0.165 (0.106)	0.157 (0.104)	0.0529 (0.119)	0.0529 (0.119)
2 Year College Degree	0.285* (0.112)	0.284* (0.112)	0.324* (0.159)	0.326* (0.159)	0.188 (0.149)	0.188 (0.149)
4 Year College Degree	0.289* (0.117)	0.288* (0.117)	0.311 (0.171)	0.310 (0.171)	0.154 (0.162)	0.154 (0.162)
Masters Degree	0.422** (0.148)	0.423** (0.148)	0.464* (0.221)	0.467* (0.222)	0.219 (0.202)	0.219 (0.203)
Doctoral Degree	0.594** (0.188)	0.594** (0.188)	0.641* (0.284)	0.649* (0.285)	0.438 (0.248)	0.438 (0.248)
Industry Controls	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1129	1129	576	576	553	553
Adjusted R^2	0.221	0.221	0.221	0.220	0.243	0.242

Notes: Industry controls include which 1 of the 16 primary ACS industry codes the respondent's primary employer was classified as in 2012. Location controls include: whether the respondent lived in a rural or urban area; whether or not they live in an MSA, and if so if it were in the central city or not; and which of the 4 primary census regions the respondents lived in at the time of the survey.

Robust standard errors in parentheses

– * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.119: Average Hourly Wage and Gender Expectations Interaction

$Y = \ln(Avg.HourlyWage_{i,2012})$								
	1997 Sub-sample				2001 Sub-sample			
	OLS	$\tau = .25$	$\tau = .5$ (Median)	$\tau = .75$	OLS	$\tau = .25$	$\tau = .5$ (Median)	$\tau = .75$
College Expectation 1997	0.00126 (0.000838)	0.000947 (0.000580)	0.000931* (0.000421)	0.00209*** (0.000549)				
College Expectation 2001					0.00192* (0.000971)	0.00286*** (0.000781)	0.00236*** (0.000668)	0.00169 (0.000983)
Male x Expectations	-0.000650 (0.00111)	-0.000696 (0.000849)	-0.00123 (0.000737)	-0.00150 (0.000870)	0.000173 (0.00118)	-0.00158 (0.000949)	-0.00124 (0.000908)	0.000137 (0.00112)
Male	0.200* (0.0941)	0.171* (0.0668)	0.265*** (0.0577)	0.323*** (0.0672)	0.179 (0.104)	0.272*** (0.0769)	0.269*** (0.0775)	0.244* (0.0983)
White	0.0592 (0.0508)	0.0338 (0.0405)	0.0288 (0.0355)	0.0189 (0.0427)	0.0153 (0.0658)	-0.0222 (0.0569)	-0.00616 (0.0467)	0.0668 (0.0544)
Black	-0.112* (0.0522)	-0.0278 (0.0361)	-0.0503 (0.0352)	0.00471 (0.0417)	-0.0474 (0.0679)	-0.117* (0.0551)	0.00247 (0.0437)	0.115* (0.0513)
Experience	0.0783* (0.0394)	0.0745* (0.0312)	-0.0103 (0.0329)	0.0142 (0.0264)	0.0416 (0.0341)	0.0714 (0.0385)	0.0506 (0.0315)	0.0461 (0.0247)
Experience ²	-0.000960 (0.00124)	-0.000787 (0.00105)	0.000603 (0.00108)	0.000573 (0.000533)	0.0000639 (0.00155)	-0.00154 (0.00192)	0.0000537 (0.00145)	0.0000674 (0.00116)
Married	0.162*** (0.0328)	0.152*** (0.0284)	0.145*** (0.0261)	0.130*** (0.0318)	0.233*** (0.0467)	0.272*** (0.0340)	0.168*** (0.0343)	0.178*** (0.0390)
ASVAB Percentile	0.00288*** (0.000840)	0.00313*** (0.000626)	0.00334*** (0.000629)	0.00399*** (0.000715)	0.00278* (0.00109)	0.00251** (0.000817)	0.00266*** (0.000762)	0.00376*** (0.000746)
Years of Schooling	0.0561 (0.0290)	0.0643** (0.0233)	0.0127 (0.0216)	0.0304 (0.0223)	0.0497* (0.0206)	0.0357 (0.0190)	0.0600** (0.0219)	0.0269 (0.0170)
High School Degree	0.169** (0.0551)	0.156** (0.0515)	0.150*** (0.0378)	0.134** (0.0491)	0.158 (0.0814)	0.183* (0.0734)	0.104 (0.0636)	0.135 (0.0739)
2 Year College Degree	0.314*** (0.0829)	0.246** (0.0748)	0.292*** (0.0613)	0.381*** (0.0657)	0.290** (0.112)	0.337** (0.108)	0.197* (0.0850)	0.343*** (0.0929)
4 Year College Degree	0.410*** (0.0852)	0.388*** (0.0721)	0.375*** (0.0705)	0.380*** (0.0711)	0.307** (0.117)	0.385*** (0.112)	0.290** (0.0928)	0.445*** (0.0993)
Masters Degree	0.447*** (0.121)	0.508*** (0.0916)	0.377*** (0.0960)	0.410*** (0.0922)	0.440** (0.147)	0.494*** (0.121)	0.381** (0.119)	0.597*** (0.119)
Doctoral Degree	0.896*** (0.166)	0.777*** (0.148)	0.896*** (0.177)	0.943*** (0.105)	0.611** (0.188)	0.676* (0.311)	0.625*** (0.173)	0.902*** (0.228)
Industry Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1860	1860	1860	1860	1129	1129	1129	1129
Adjusted R^2	0.193				0.220			
Pseudo R^2		0.1463	0.1444	0.1385		0.1729	0.1718	0.1595

Notes: Industry controls include which 1 of the 16 primary ACS industry codes the respondent's primary employer was classified as in 2012. Location controls include: whether the respondent lived in a rural or urban area; whether or not they live in an MSA, and if so if it were in the central city or not; and which of the 4 primary census regions the respondents lived in at the time of the survey.

Robust standard errors in parentheses

- * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.120: Average Hourly Wage and Race Expectations Interaction

$Y = \text{Ln}(\text{Avg.HourlyWage}_{i,2012})$						
	1997 Sub-sample: Full, Men, Women			2001 Sub-sample: Full, Men, Women		
	(1)	(2)	(3)	(4)	(5)	(6)
College Expectation 1997	-0.0000993 (0.00112)	0.00112 (0.00146)	-0.00180 (0.00178)			
College Expectation 2001				0.00189 (0.00129)	0.00340 (0.00181)	0.00142 (0.00180)
White x Expectations	0.000606 (0.00132)	-0.00127 (0.00171)	0.00355 (0.00210)	0.000681 (0.00151)	-0.00111 (0.00198)	0.00235 (0.00230)
Black x Expectations	0.00269 (0.00168)	0.00281 (0.00266)	0.00295 (0.00217)	.000738 (0.00171)	-0.00272 (0.00247)	-0.000603 (0.00239)
White	0.0183 (0.113)	0.115 (0.141)	-0.172 (0.185)	-0.0329 (0.135)	0.0591 (0.163)	-0.170 (0.214)
Black	-0.316* (0.135)	-0.405* (0.205)	-0.239 (0.179)	0.00598 (0.149)	-0.0119 (0.208)	0.111 (0.208)
Male	0.149*** (0.0359)			0.191*** (0.0460)		
ASVAB Percentile	0.00285*** (0.000838)	0.00221 (0.00118)	0.00395** (0.00121)	0.00278* (0.00109)	0.00101 (0.00161)	0.00461** (0.00145)
Experience	0.0784 (0.0414)	0.0636 (0.0618)	0.0916 (0.0604)	0.0431 (0.0368)	0.119* (0.0488)	-0.0320 (0.0425)
Experience ²	-0.000891 (0.00125)	0.000822 (0.00166)	-0.00247 (0.00199)	-0.00000644 (0.00154)	-0.00283 (0.00192)	0.00305 (0.00170)
Married	0.159*** (0.0327)	0.213*** (0.0496)	0.106* (0.0440)	0.234*** (0.0471)	0.331*** (0.0729)	0.124* (0.0611)
Years of Schooling	0.0564 (0.0291)	0.107* (0.0452)	0.0125 (0.0368)	0.0492* (0.0206)	0.0710* (0.0318)	0.0414 (0.0264)
High School Degree	0.168** (0.0551)	0.167* (0.0824)	0.137 (0.0715)	0.160* (0.0810)	0.171 (0.105)	0.0677 (0.116)
2 Year College Degree	0.312*** (0.0827)	0.232 (0.124)	0.384*** (0.109)	0.292** (0.112)	0.330* (0.158)	0.200 (0.149)
4 Year College Degree	0.409*** (0.0847)	0.306* (0.125)	0.478*** (0.118)	0.303** (0.116)	0.331 (0.171)	0.157 (0.159)
Masters Degree	0.446*** (0.121)	0.0654 (0.232)	0.680*** (0.149)	0.435** (0.147)	0.477* (0.221)	0.218 (0.200)
Doctoral Degree	0.894*** (0.164)	0.559* (0.269)	1.070*** (0.197)	0.609** (0.188)	0.660* (0.285)	0.453 (0.247)
Industry Controls	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1860	959	901	1129	576	553
Adjusted R^2	0.194	0.156	0.242	0.220	0.220	0.245

Notes: Industry controls include which 1 of the 16 primary ACS industry codes the respondent's primary employer was classified as in 2012 . Location controls include: whether the respondent lived in a rural or urban area; whether or not they live in an MSA, and if so if it were in the central city or not; and which of the 4 primary census regions the respondents lived in at the time of the survey.

Robust standard errors in parentheses

– * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Bibliography

- Allen, I. E. and J. Seaman (2010). Class differences: Online education in the united states, 2010. *Sloan Consortium (NJ1)*.
- Altonji, J. (1993). The demand for and return to education when education outcomes are uncertain. *Journal of Labor Economics* 11(1), 48–83.
- Altonji, J. G. and C. R. Pierret (1996). Employer learning and the signaling value of education. Technical report, National Bureau of Economic Research.
- Angrist, J. D. and A. B. Keueger (1991). Does compulsory school attendance affect schooling and earnings? *The Quarterly Journal of Economics* 106(4), 979–1014.
- Antonovics, K. and R. Town (2004). Are all the good men married? uncovering the sources of the marital wage premium. *The American Economic Review* 94(2), 317–321.
- Arcidiacono, P. (2004). Ability sorting and the returns to college major. *Journal of Econometrics* 121(1), 343–375.
- Arcidiacono, P., V. J. Hotz, and S. Kang (2012). Modeling college major choices using elicited measures of expectations and counterfactuals. *Journal of Econometrics* 166(1), 3–16.
- Bailey, M. J. and S. Dynarski (2011). Gains and gaps: Changing inequality in us college entry and completion. Technical report, National Bureau of Economic Research.
- Bamberger, G. (1987). *Occupational choice: the role of undergraduate education*. Ph. D. thesis, Univ. of Chicago, Graduate School of Business.

- Becker, G. S. (1994). Human capital revisited. In *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education (3rd Edition)*, pp. 15–28. The University of Chicago Press.
- Bernheim, B. D. (1990). How do the elderly form expectations? an analysis of responses to new information. In *Issues in the economics of aging*, pp. 259–286. University of Chicago Press.
- Bettinger, E. P. and B. T. Long (2009). Addressing the needs of underprepared students in higher education does college remediation work? *Journal of Human resources* 44(3), 736–771.
- Black, D., S. Sanders, and L. Taylor (2011). Measurement of higher education in the census and current population survey. *Journal of the American Statistical Association*.
- Blomquist, G. C., P. A. Coomes, C. Jepsen, B. C. Koford, and K. R. Troske (2014). Estimating the social value of higher education: willingness to pay for community and technical colleges. *Journal of Benefit-Cost Analysis* 5(1), 3–41.
- Bollinger, C. R. (2003). Measurement error in human capital and the black-white wage gap. *Review of Economics and Statistics* 85(3), 578–585.
- Bureau of Labor Statistics, U.S. Department of Labor (2013). *National Longitudinal Survey of Youth 1997 cohort, 1997-2011 (rounds 1-15)*. Columbus, OH:: Produced by the National Opinion Research Center, the University of Chicago and distributed by the Center for Human Resource Research, The Ohio State University.
- Card, D. (1999). The causal effect of education on earnings. *Handbook of labor economics* 3, 1801–1863.
- Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica* 69(5), 1127–1160.

- CollegeBoard (2016). Trends in college pricing 2016. Technical report, www.collegeboard.org.
- Cornwell, C. and P. Rupert (1997). Unobservable individual effects, marriage and the earnings of young men. *Economic Inquiry* 35(2), 285–294.
- Cutler, D. M. and A. Lleras-Muney (2006). Education and health: evaluating theories and evidence. Technical report, National Bureau of Economic Research.
- Davies, M. and D. B. Kandel (1981). Parental and peer influences on adolescents' educational plans: Some further evidence. *American Journal of Sociology* 87(2), 363–387.
- Department of Education, N. C. f. E. S. (2013). *Digest of Education Statistics: 2012*. http://nces.ed.gov/programs/digest/d12/ch_3.asp.
- Dominitz, J. and C. F. Manski (1996). Eliciting student expectations of the returns to schooling. *Journal of Human Resources* 31, 1–26.
- Dominitz, J. and C. F. Manski (1997). Using expectations data to study subjective income expectations. *Journal of the American Statistical Association* 92(439), 855–867.
- Dynarski, S. (2000). Hope for whom? financial aid for the middle class and its impact on college attendance. Technical report, National bureau of economic research.
- Dynarski, S. M. (1999). Does aid matter? measuring the effect of student aid on college attendance and completion. Technical report, National bureau of economic research.
- Fischhoff, B., A. M. Parker, W. B. de Bruin, J. Downs, C. Palmgren, R. Dawes, and C. F. Manski (2000). Teen expectations for significant life events. *Public Opinion Quarterly*, 189–205.

- Forbes (2016). America's top colleges: 2016 ranking.
- Ginther, D. K. and M. Zavodny (2001). Is the male marriage premium due to selection? the effect of shotgun weddings on the return to marriage. *Journal of Population Economics* 14(2), 313–328.
- Greenstone, M. and A. Looney (2012). Regardless of the cost, college still matters. Technical report, The Hamilton Project.
- Grossman, M. (2006). Education and nonmarket outcomes. *Handbook of the Economics of Education* 1, 577–633.
- Hanson, S. L. (1994). Lost talent: Unrealized educational aspirations and expectations among us youths. *Sociology of Education*, 159–183.
- Heckman, J. J., L. J. Lochner, and P. E. Todd (2008). Earnings functions and rates of return. Technical report, National Bureau of Economic Research.
- Hilmer, M. J. (1998). Post-secondary fees and the decision to attend a university or a community college. *Journal of Public Economics* 67(3), 329–348.
- Hoelter, J. W. (1982). Segregation and rationality in black status aspiration processes. *Sociology of Education*, 31–39.
- Hossler, D. and K. S. Gallagher (1987). Studying student college choice: A three-phase model and the implications for policymakers. *College and university* 62(3), 207–21.
- Hossler, D., J. Schmit, and N. Vesper (1999). *Going to college: How social, economic, and educational factors influence the decisions students make*. JHU Press.
- Hossler, D. and F. K. Stage (1992). Family and high school experience influences on the postsecondary educational plans of ninth-grade students. *American Educational Research Journal* 29(2), 425–451.

- Hoxby, C., S. Turner, et al. (2013). Expanding college opportunities for high-achieving, low income students. *Stanford Institute for Economic Policy Research Discussion Paper* (12-014).
- Hoxby, C. M. and C. Avery (2012). The missing "one-offs": The hidden supply of high-achieving, low income students. Technical report, National Bureau of Economic Research.
- Hunter, M. G. (2017a). *Adolescent Econometricians: Smarter than You Might Think*. Ph. D. thesis, University of Kentucky.
- Hunter, M. G. (2017b). *The Formation and Updating of Post-secondary Educational Expectations as Student's Transition from High School*. Ph. D. thesis, University of Kentucky.
- Hunter, M. G. (2017c). *TStudent Predictions and Alignment of Future College Success*. Ph. D. thesis, University of Kentucky.
- Hurd, M. D. and K. McGarry (2002a). The predictive validity of subjective probabilities of survival. *The Economic Journal* 112(482), 966–985.
- Hurd, M. D. and K. McGarry (2002b). The predictive validity of subjective probabilities of survival. *The Economic Journal* 112(482), 966–985.
- Jacob, B. A. and T. Wilder (2011). *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances*, Chapter 7: Educational Expectations and Attainment. Russell Sage Foundation.
- John, E. P. S., M. B. Paulsen, and J. B. Starkey (1996). The nexus between college choice and persistence. *Research in Higher Education* 37(2), 175–220.
- Kane, T. J. and C. E. Rouse (1995). Labor-market returns to two-and four-year college. *The American Economic Review*, 600–614.

- Kao, G. and M. Tienda (1998). Educational aspirations of minority youth. *American journal of education* 106(3), 349–384.
- Kiuru, N., K. Aunola, J. Vuori, and J.-E. Nurmi (2007). The role of peer groups in adolescents’s educational expectations and adjustment. *Journal of Youth and Adolescence* 36(8), 995–1009.
- Kroeger, T., C. Tanyell, and E. Gould (2016). The class of 2016: The labor market is still far from ideal for young graduates. Technical report, Economic Policy Institute.
- Krueger, A. and O. Ashenfelter (1994). Estimates of the economic return to schooling from a new sample of twins. *The American Economic Review* 85(5), 1157–1173.
- Leslie, L. L. and P. T. Brinkman (1988). *The Economic Value of Higher Education*. American Council on Education/Macmillan Series on Higher Education. ERIC.
- Light, A. (2004). Gender differences in the marriage and cohabitation income premium. *Demography* 41(2), 263–284.
- Lochner, L. (2007). Individual perceptions of the criminal justice system. *The American Economic Review* 97(1), 444–460.
- Lochner, L. and E. Moretti (2001). The effect of education on crime: Evidence from prison inmates, arrests, and self-reports. Technical report, National Bureau of Economic Research.
- Lundberg, S. (1988). Labor supply of husbands and wives: A simultaneous equations approach. *The Review of Economics and Statistics*, 224–235.
- Lundberg, S. J., R. A. Pollak, and T. J. Wales (1997). Do husbands and wives pool their resources? evidence from the united kingdom child benefit. *Journal of Human resources*, 463–480.

- Machlup, F. (1946). Marginal analysis and empirical research. *The American Economic Review* 36(4), 519–554.
- MacLeod, J. (2008). *Ain't no makin'it: Aspirations and attainment in a low-income neighborhood*. Westview Press.
- Malumud, O. (2005). *Breadth vs. depth: The effect of academic specialization on labor market outcomes*. Harris School of Public Policy, University of Chicago.
- Manski, C. F. (1989). Schooling as experimentation: a reappraisal of the postsecondary dropout phenomenon. *Economics of Education review* 8(4), 305–312.
- Manski, C. F. (1993). Adolescent econometricians: How do youth infer the returns to schooling? In *Studies of supply and demand in higher education*, pp. 43–60. University of Chicago Press.
- Manski, C. F. (2004). Measuring expectations. *Econometrica* 72(5), 1329–1376.
- Manski, C. F. and D. A. Wise (1983). *College choice in America*. Harvard University Press.
- Mincer, J. (1974). Schooling, experience, and earnings.
- Mroz, T. A. (1987). The sensitivity of an empirical model of married women's hours of work to economic and statistical assumptions. *Econometrica: Journal of the Econometric Society*, 765–799.
- Neal, D. A. and W. R. Johnson (1996). The role of premarket factors in black-white wage differences. *Journal of political Economy* 104(5), 869–895.
- Oreopoulos, P. and U. Petronijevic (2013). Making college worth it: A review of the returns to higher education. *The Future of Children* 23(1), 41–65.
- Pencavel, J. (1986). Labor supply of men: a survey. *Handbook of labor economics* 1,

3–102.

Perna, L. W. (2000). Differences in the decision to attend college among african americans, hispanics, and whites. *Journal of Higher Education*, 117–141.

Polachek, S. W. et al. (2008). Earnings over the life cycle: The mincer earnings function and its applications. *Foundations and Trends® in Microeconomics* 4(3), 165–272.

Polachek, S. W. and W. S. Siebert (1993). *The economics of earnings*. Cambridge University Press.

Sewell, W. H., A. O. Haller, and G. W. Ohlendorf (1970). The educational and early occupational status attainment process: Replication and revision. *American Sociological Review*, 1014–1027.

Shapiro, D., A. Dundar, P. Wakhungu, X. Yuan, A. Nathan, and Y. Hwang (2015). Completing college: A national view of student attainment rates – fall 2009 cohort. Technical report, National Student Clearinghouse Research Center.

Smith, V. K., D. H. Taylor, and F. A. Sloan (2001). Longevity expectations and death: Can people predict their own demise? *The American Economic Review* 91(4), 1126–1134.

St. John, E. P., M. B. Paulsen, and D. F. Carter (2005). Diversity, college costs, and postsecondary opportunity: An examination of the financial nexus between college choice and persistence for african americans and whites. *The Journal of Higher Education* 76(5), 545–569.

Stage, F. K. and D. Hossler (1989). Differences in family influences on college attendance plans for male and female ninth graders. *Research in Higher Education* 30(3), 301–315.

- Stewart, E. B., E. A. Stewart, and R. L. Simons (2007). The effect of neighborhood context on the college aspirations of african american adolescents. *American Educational Research Journal* 44(4), 896–919.
- Strange, K. M. (2012). An empirical investigation of the option value of college enrollment. *American Economic Journal: Applied Economics* 4(1), 49–84.
- Teachman, J. D. and K. Paasch (1998). The family and educational aspirations. *Journal of Marriage and the Family*, 704–714.
- Terenzini, P. T., A. F. Cabrera, and E. M. Bernal (2001). Swimming against the tide: The poor in american higher education. research report no. 2001-1. *College Entrance Examination Board*.
- Tyler, J. H., R. J. Murnane, and J. B. Willett (2000). Estimating the labor market signaling value of the ged. *Quarterly Journal of Economics*, 431–468.
- U.S. Department of Education (2017). *National Longitudinal Survey of Youth 1997 cohort, 1997-2011 (rounds 1-15)*. Washington, DC: Office of Postsecondary Education.
- Vedder, R., C. Denhart, and J. Robe (2013). Why are recent college graduates underemployed? university enrollments and labor-market realities. *Center for College Affordability and Productivity (NJ1)*.
- Willis, R. J. (1986). Wage determinants: A survey and reinterpretation of human capital earnings functions. *Handbook of labor economics* 1, 525–602.
- Zafar, B. (2011). How do college students form expectations? *Journal of Labor Economics* 29(2), 301–348.
- Zafar, B. (2013). College major choice and the gender gap. *Journal of Human Resources* 48(3), 545–595.

Education

M.S., Economics, University of Kentucky, Lexington, KY (December 2013)

B.S., Financial Economics, Centre College, Danville, KY (May 2011)

Teaching Positions

Lecturer, University of Arizona (Fall 2017 – Present)

Primary Instructor, Transylvania University (Summer 2016)

Primary Instructor, University of Kentucky (Fall 2013 - Summer 2015)

Teaching Assistant, University of Kentucky (Fall 2012 – Spring 2014)

Awards and Honors

Outstanding Teaching Assistant Award from the University of Kentucky Economics Department (Academic Year 2015 – 2016)

University of Kentucky Lockett Fellowship, (Academic Year 2016 – 2017)

University of Kentucky Lockett Fellowship, (Academic Year 2015 – 2016)

University of Kentucky Gatton Doctoral Fellowship, (Academic Year 2013 – 2014, 2014-2015)

University of Kentucky, Max Steckler Fellowship (Academic Year 2012-2013)

Research

“Adolescent Econometricians: Smarter than You Might Think.” working paper

“How Students Form and Update their Educational Expectations as they Transition from High School” working paper

“Modeling Educational Expectations and College Choice” working paper

Research and Professional Experience

Research Assistant, University of Kentucky Center for Business and Economic Research, (June 2015-Present)

Research Assistant, Centre College Department of Economics, (June 2010 - August 2010)

Teaching Enhancement Activities

AEA Conference on Teaching and Research in Economic Education (discussant), Atlanta, GA (June 2016)

ECO 201 Principles of Microeconomics Substitute for 450 Student Lecture, University of Kentucky (September 2015)

New Classroom Technology Training (participant), University of Kentucky (September 2015)

ECO 391 Economics and Business Statistics Student Survey Coordinator (Data Manager), Department of Economics, University of Kentucky (January 2015-May 2015)

University of Kentucky 20th Annual Economics Teaching Workshop (presenter), Lexington, KY (April 2015)

Turning Point Clicker Training Workshop (participant), University of Kentucky (August 2014)

University of Kentucky Annual Economics Teaching Workshop (participant), Lexington, KY (April 2013, 2014, and 2016)

Teaching Assistant Orientation with Microteaching (participant), University of Kentucky (August 2012)

Professional Activities

Editor: *Instructor Material Revisionist*, Bedford, Freeman, & Worth Publishing Group, LLC (Summer 2015)

Member: *Textbook Selection Committee*, Economics Department, University of Kentucky (Spring 2015)

Conferences

Association for Education Finance and Policy 41st Annual Meeting (presenter), Denver, CO (March 2016)

Southern Economic Association 85th Annual Meeting (discussant), New Orleans, LA (November 2015)

Kentucky Economic Association Annual Meeting (presenter, discussant), Frankfort, KY (October 2015 and 2016)

Federal Reserve Bank of Cleveland and the University of Kentucky Labor Market Workshop (participant), Multiple Locations (February 2015 and 2016)

Kentucky Annual Economic Outlook Conference (participant) , Lexington, KY (February 2015 and 2016)

Other Experience

Contract Auditor, *Dealer's Financial Services*, Lexington, KY (February 2012 – August 2012)

Financial Service Representative, *First Investors Corporation*, Lexington, KY (July 2011 – December 2011)